Optimal Transport and Equilibrium Problems in Mathematical Finance

Xiaowei Tan

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

Abstract

Optimal Transport and Equilibrium Problems in Mathematical Finance

Xiaowei Tan

The thesis consists of three independent topics, each of which is discussed in an individual chapter.

The first chapter considers a multiperiod optimal transport problem where distributions $\mu_0, \dots, \mu_n$ are prescribed and a transport corresponds to a scalar martingale X with marginals $X_t \sim \mu_t$. We introduce particular couplings called left-monotone transports; they are characterized equivalently by a no-crossing property of their support, as simultaneous optimizers for a class of bivariate transport cost functions with a Spence–Mirrlees property, and by an order-theoretic minimality property. Left-monotone transports are unique if μ_0 is atomless, but not in general. In the one-period case $n = 1$, these transports reduce to the Left-Curtain coupling of Beiglböck and Juillet. In the multiperiod case, the bivariate marginals for dates $(0, t)$ are of Left-Curtain type, if and only if $\mu_0, \dots, \mu_n$ have a specific order property. The general analysis of the transport problem also gives rise to a strong duality result and a description of its polar sets. Finally, we study a variant where the intermediate marginals $\mu_1, \dots, \mu_{n-1}$ are not prescribed.

The second chapter studies the convergence of Nash equilibria in a game of optimal stopping. If the associated mean field game has a unique equilibrium, any sequence of n-player equilibria converges to it as $n \to \infty$. However, both the finite and infinite player versions of the game often admit multiple equilibria. We show that mean field equilibria satisfying a transversality condition

are limit points of n-player equilibria, but we also exhibit a remarkable class of mean field equilibria that are not limits, thus questioning their interpretation as "large n" equilibria.

The third chapter studies the equilibrium price of an asset that is traded in continuous time between N agents who have heterogeneous beliefs about the state process underlying the asset's payoff. We propose a tractable model where agents maximize expected returns under quadratic costs on inventories and trading rates. The unique equilibrium price is characterized by a weakly coupled system of linear parabolic equations which shows that holding and liquidity costs play dual roles. We derive the leading-order asymptotics for small transaction and holding costs which give further insight into the equilibrium and the consequences of illiquidity.

Contents

List of Figures

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor Marcel Nutz for his inspiring guidance and invaluable encouragement over the years. It's also my pleasure to thank Julien Dubedat, Ioannis Karatzas, Dylan Possamai, Philip Protter for serving on my dissertation committee.

I am indebted to my collaborators Johannes Muhle-Karbe, Jaime San Martin, Florian Stebegg for their numerous contributions to our papers. During my time as a doctoral student, I also benefitted a lot from discussions with Zhi Li, Minghan Yan, Yang Kang, Fengpei Li, Yuchong Zhang, Cameron Bruggeman, Praveen Kolli, Yinghui Wang. Furthermore, I am grateful to my previous roommates Shizhang Li, Zhechi Cheng, and my best friend Run Wang for all the great memories we had together.

Last but not the least, I am deeply thankful to my family and friends for their selfless love and generous support in every aspect of my life.

To my parents Shaobin Tan and Zhiwei Chen,

my younger brother Jiawei Tan

and

my fiancée Natalie Wang

Chapter 1

Multiperiod Martingale Transport

1.1 Introduction

Let $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ be a vector of probability measures μ_t on the real line. A measure P on $\mathbb{R}^{n+1}$ whose marginals are given by $\boldsymbol{\mu}$ is called a coupling (or transport) of $\boldsymbol{\mu}$, and the set of all such measures is denoted by $\Pi(\boldsymbol{\mu})$. We shall be interested in couplings P that are martingales; that is, the identity $X = (X_0, \ldots, X_n)$ on $\mathbb{R}^{n+1}$ is a martingale under P. Hence, we will assume that all marginals have a finite first moment and denote by $\mathcal{M}(\boldsymbol{\mu})$ the set of martingale couplings. A classical result of Strassen [104] shows that $\mathcal{M}(\boldsymbol{\mu})$ is nonempty if and only if the marginals are in convex order, denoted by $\mu_{t-1} \leq_c \mu_t$ and defined by the requirement that $\mu_{t-1}(\phi) \leq \mu_t(\phi)$ for any convex function ϕ, where $\mu(\phi) := \int \phi \, d\mu$.

The first goal of this paper is to introduce and study a family of "canonical" couplings $P \in \mathcal{M}(\mu)$ that we call left-monotone. These couplings specialize to the Left-Curtain coupling of [16] in the one-step case $n = 1$ and share, broadly speaking, several properties reminiscent of the Hoeffding–Fréchet coupling of classical optimal transport. Indeed, left-monotone couplings will be characterized by order-theoretic minimality properties, as simultaneous optimal transports for certain classes of reward (or cost) functions, and through no-crossing conditions on their supports.

The second goal is to develop a strong duality theory for multiperiod martingale optimal transport, along the lines of [18] for the one-period martingale case and [78] for the classical optimal transport problem. That is, we introduce a suitable dual optimization problem and show the absence of a duality gap as well as the existence of dual optimizers for general transport reward (or cost) functions. The duality result is a crucial tool for the study of the left-monotone couplings.

We also develop similar results for a variant of our problem where the intermediate marginals $\mu_1, \dots, \mu_{n-1}$ are not prescribed (Section 1.9), but we shall focus on the full marginal case for the purpose of the Introduction.

1.1.1 Left-Monotone Transports

For the sake of orientation, let us first state the main result and then explain the terminology contained therein. The following is a streamlined version—the results in the body of the paper are stronger in some technical aspects.

Theorem 1.1.1. *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order and $P \in \mathcal{M}(\boldsymbol{\mu})$ a martingale transport between these marginals. The following are equivalent:*

(i) P is a simultaneous optimal transport for $f(X_0, X_t)$, $1 \leq t \leq n$ whenever $f : \mathbb{R}^2 \to \mathbb{R}$ is a smooth second-order Spence–Mirrlees function.

(ii) P is concentrated on a left-monotone set $\Gamma \subseteq \mathbb{R}^{n+1}$.

(iii) P transports $\mu_0|_{(-\infty,a]}$ to the obstructed shadow $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{(-\infty,a]})$ in step t, for all $1 \leq t \leq n$ and $a \in \mathbb{R}$.

There exists $P \in \mathcal{M}(\boldsymbol{\mu})$ satisfying (i)–(iii), and any such P is called a left-monotone transport. If μ_0 is atomless, then P is unique.

Let us now discuss the items in the theorem.

(i) Optimal Transport. This property characterizes P as a simultaneous optimal transport. Given a function $f : \mathbb{R}^{n+1} \to \mathbb{R}$, we may consider the martingale optimal transport problem with reward f (or cost $-f$),

$$\mathbf{S}_{\boldsymbol{\mu}}(f) = \sup_{P\in\mathcal{M}(\boldsymbol{\mu})} P(f); \tag{1.1.1}$$

recall that $P(f) = \mathbb{E}^P[f(X_0,\dots,X_n)]$. A Lipschitz function $f \in C^{1,2}(\mathbb{R}^2;\mathbb{R})$ is called a *smooth second-order Spence–Mirrlees* function if it satisfies the cross-derivative condition $f_{xyy} > 0$; this has also been called the martingale Spence–Mirrlees condition in analogy to the classical Spence–Mirrlees condition $f_{xy} > 0$. Given such a function of two variables and $1 \le t \le n$, we may consider the n-step martingale optimal transport problem with reward $f(X_0, X_t)$. Characterization (i) states that a left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ is an optimizer simultaneously for the n transport problems $f(X_0, X_t)$, $1 \le t \le n$, for some (and then all) smooth second-order Spence–Mirrlees functions f.

In the one-step case, a corresponding result holds for the Left-Curtain coupling [16]; here the simultaneous optimization becomes a single one. In view of the characterization in (i), an immediate consequence is that if there exists $P \in \mathcal{M}(\boldsymbol{\mu})$ such that all bivariate projections $P_{0t} = P \circ (X_0, X_t)^{-1} \in \mathcal{M}(\mu_0, \mu_t)$ are of Left-Curtain type, then P is left-monotone. However, such a transport does not exist unless the marginals satisfy a very specific condition (see Proposition 1.6.9), and in general the bivariate projections of a left-monotone transport are *not* of Left-Curtain type.

(ii) Geometry. The second item characterizes P through a geometric property of its support. A set $\Gamma \subseteq \mathbb{R}^{n+1}$ will be called *left-monotone* if it has the following no-crossing property for all $1 \le t \le n$: Let $\boldsymbol{x} = (x_0, \dots, x_{t-1})$, $\boldsymbol{x}' = (x'_0, \dots, x'_{t-1}) \in \mathbb{R}^t$ and

$$y^-, y^+, y' \in \mathbb{R} \quad \text{with} \quad y^- < y^+$$

be such that $(\boldsymbol{x}, y^+), (\boldsymbol{x}, y^-), (\boldsymbol{x}', y')$ are in the projection of Γ to the first $t+1$ coordinates. Then,

$$y' \notin (y^-, y^+) \quad \text{whenever} \quad x_0 < x'_0.$$

That is, if we consider two paths in Γ starting at x_0 and coinciding up to $t-1$, and a third path starting at x'_0 to the right of x_0, then at time t the third path cannot step in-between the first two—this is illustrated in Figure 1.1. Item (ii) states that a left-monotone transport

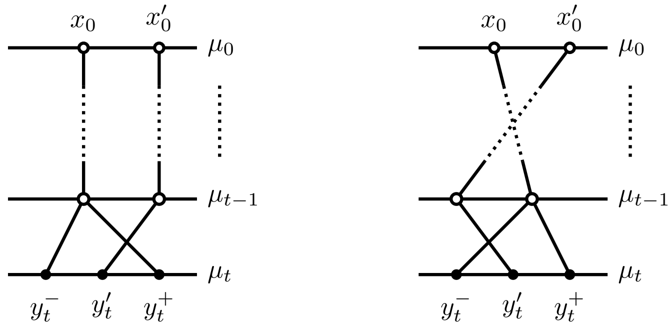


Figure 1.1: Two examples of forbidden configurations in left-monotone sets.

$P \in \mathcal{M}(\boldsymbol{\mu})$ can be characterized by the fact that it is concentrated on a left-monotone set Γ. (In Theorem 1.7.16 we shall state a stronger result: we can find a left-monotone set that carries all left-monotone transports at once.)

In the one-step case $n = 1$, left-monotonicity coincides with the Left-Curtain property of [16]. However, we emphasize that for $t > 1$, our no-crossing condition differs from the Left-Curtain property of the bivariate projection $(X_0, X_t)(\Gamma)$ as the latter would not contain the restriction that the first two paths have to coincide up to $t-1$ (see also Example 1.6.10). This corresponds to the mentioned fact that the bivariate marginal P_{0t} need not be of Left-Curtain type. On the other hand, the geometry of the projection $(X_{t-1}, X_t)(\Gamma)$ is also quite different from the Left-Curtain one, as our condition may rule out third paths crossing from the right *and* left at $t-1$, depending on the starting point x_0' rather than the location of x_{t-1}'.

(iii) Convex Ordering. This property characterizes left-monotone transports in an order-theoretic way and will be used in the existence proof. To explain the idea, suppose that μ_0 consists of finitely many atoms at $x_1, \dots, x_N \in \mathbb{R}$. Then, for any fixed t, a coupling of μ_0 and μ_t can be defined by specifying a "destination" measure for each atom. We consider all chains[1] $\mu_0|_{x_i} \leq_c \theta_1 \leq_c \cdots \leq_c \theta_t$ of measures θ_s in convex order that satisfy the marginals constraints $\theta_s \leq \mu_s$ for $s \leq t$. Of these chains, keep only the terminal measures θ_t and compare them according to the convex order. The *obstructed shadow of* $\mu_0|_{x_1}$ *in* μ_t *through*

[1]Here $\mu_0|_{x_i}$ denotes a Dirac measure of mass $\mu_0(\{x_i\})$ at x_i.

$\mu_1, \dots, \mu_{t-1}$, denoted $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{x_i})$, is defined as the unique least element[2] among the θ_t. A particular coupling of μ_0 and μ_t is the one that successively maps the atoms $\mu_0|_{x_i}$ to their obstructed shadows in the remainder of μ_t, starting with the left-most atom x_i and continuing from left to right. In the case of general measures, we consider the restrictions $\mu_0|_{(-\infty,a]}$ instead of successively mapping the atoms. Characterization (iii) then states that a left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ maps $\mu_0|_{(-\infty,a]}$ to its obstructed shadow at date t for all $1 \leq t \leq n$ and $a \in \mathbb{R}$. This shows in particular that the bivariate projections $P_{0t} = P \circ (X_0, X_t)^{-1}$ of a left-monotone coupling are uniquely determined. In the body of the text, we shall also give an alternative definition of the obstructed shadow by iterating unobstructed shadows through the marginals up to date t; see Section 1.6.

The above specializes to the construction of [16] for the one-step case, which corresponds to the situation of $t = 1$ where there are no intermediate marginals obstructing the shadow. When $t > 1$, the obstruction by the intermediate marginals once again entails that P_{0t} need not be of Left-Curtain type. More precisely, Characterization (iii) gives rise to a sharp criterion (Proposition 1.6.9) on the marginals $\boldsymbol{\mu}$, describing exactly when this coincidence arises.

(Non-)Uniqueness. We have seen above that for a left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ the bivariate projections P_{0t}, $1 \leq t \leq n$ are uniquely determined. In particular, for $n = 1$, we recover the result of [16] that the left-monotone coupling is unique. For $n > 1$, the situation turns out to be quite different depending on the nature of the first marginal. On the one extreme, we shall see that when μ_0 is atomless, there is a unique left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$. Moreover, P has a degenerate structure reminiscent of Brenier's theorem: it can be disintegrated as $P = \mu_0 \otimes \kappa_1 \otimes \cdots \otimes \kappa_n$ where each one-step transport kernel κ_t is concentrated on the graphs of two functions. On the other extreme, if μ_0 is a Dirac mass, the typical case is that there are infinitely many left-monotone couplings—see Section 1.8 for a detailed discussion. We shall also show that left-monotone transports are not Markovian in general, even if uniqueness holds (Example 1.7.17).

[2] See Definition 1.6.6 and Lemma 1.6.7 for details on this construction.

1.1.2 Duality

The analysis of left-monotone transports is based on a duality result that we develop for general reward functions $f : \mathbb{R}^{n+1} \to (-\infty, \infty]$ with an integrable lower bound. Formally, the dual problem (in the sense of linear programming) for the transport problem $\mathbf{S}_{\boldsymbol{\mu}}(f) = \sup_{P \in \mathcal{M}(\boldsymbol{\mu})} P(f)$ is the minimization

$$\mathbf{I}_{\boldsymbol{\mu}}(f) := \inf_{(\boldsymbol{\phi}, H)} \sum_{t=0}^{n} \mu_t(\phi_t)$$

where the infimum is taken over vectors $\boldsymbol{\phi} = (\phi_0, \ldots, \phi_n)$ of real functions and predictable processes $H = (H_1, \ldots, H_n)$ such that

$$\sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n \geq f; \tag{1.1.2}$$

here $(H \cdot X)_n := \sum_{t=1}^{n} H_t (X_t - X_{t-1})$ is the discrete-time integral. The desired result (Theorem 1.5.2) states that there is no duality gap, i.e. $\mathbf{I}_{\boldsymbol{\mu}}(f) = \mathbf{S}_{\boldsymbol{\mu}}(f)$, and that the dual problem is attained whenever it is finite. From the analysis for the one-step case in [18] we know that this assertion fails for the above naive formulation of the dual, and requires several relaxations regarding the integrability of the functions ϕ_t and the domain $\mathcal{V} \subseteq \mathbb{R}^{n+1}$ where the inequality (1.1.2) is required. Specifically, the inequality needs to be relaxed on sets that are $\mathcal{M}(\boldsymbol{\mu})$-polar; i.e. not charged by any transport $P \in \mathcal{M}(\boldsymbol{\mu})$. These sets are characterized in Theorem 1.3.1 where we show that the $\mathcal{M}(\boldsymbol{\mu})$-polar sets are precisely the (unions of) sets which project to a two-dimensional polar set of $\mathcal{M}(\mu_{t-1}, \mu_t)$ for some $1 \leq t \leq n$.

The duality theorem gives rise to a monotonicity principle (Theorem 1.5.4) that underpins the analysis of the left-monotone couplings. Similarly to the cyclical monotonicity condition in classical transport, it allows one to study the geometry of the support of optimal transports for a given function f.

1.1.3 Background and Related Literature

The martingale optimal transport problem (1.1.1) was introduced in [13] with the dual problem as a motivation. Indeed, in financial mathematics the function f is understood as the payoff of a derivative written on the underlying X and (1.1.2) corresponds to superhedging f by statically trading in European options $\phi_t(X_t)$ and dynamically trading in the underlying according to the strategy H. The value $\mathbf{I}_{\boldsymbol{\mu}}(f)$ then corresponds to the lowest price of f for which the seller can enter a model-free hedge $(\boldsymbol{\phi}, H)$ if the marginals $X_t \sim \mu_t$ are known from option market data. In [13], it was shown (with the above, "naive" formulation of the dual problem) that there is no duality gap if f is sufficiently regular, whereas dual existence was shown to fail even in regular cases. The idea of model-free hedging as well as the connection to Skorokhod embeddings goes back to [64]; we refer to [24, 26, 41, 65, 97, 108] for further references. A specific multiperiod martingale optimal transport problem also arises in the study of the maximum maximum of a martingale given n marginals [59].

The one-step case $n = 1$ has been studied in great detail. In particular, [16] introduced the Left-Curtain coupling and pioneered numerous ideas underlying Theorem 1.1.1, [61] provided an explicit construction of that coupling, and [73] established the stability with respect to the marginals. Our duality results specialize to the ones of [18] when $n = 1$. Unsurprisingly, we shall exploit many arguments and results from these papers wherever possible. As indicated above, and as will be seen in the proofs below, the multistep case allows for a richer structure and necessitates novel ideas; for instance, the analysis of the polar sets (Theorem 1.3.1) is surprisingly involved. Other works in the one-step martingale case have studied reward functions f such as forward start straddles [66, 67] or Asian payoffs [103]. We also refer to [55, 88] for recent developments with multidimensional marginals.

One-step martingale optimal transport problems can alternately be studied as optimal Skorokhod embedding problems with marginal constraints; cf. [10, 11, 14, 15]. A multi-marginal extension [9] of [10] is in preparation at the time of writing and the authors have brought to our attention that it will offer a version of Theorem 1.1.1 in the Skorokhod picture, at least in the case where μ_0 is atomless and some further conditions are satisfied. The Skorokhod embedding problem with multi-marginal constraint was also studied in [56].

A multi-step coupling quite different from ours can be obtained by composing in a

Markovian fashion the Left-Curtain transport kernels from μ_{t-1} to μ_t, $1 \leq t \leq n$, as discussed in [61]. In [74] the continuous-time limits of such couplings for $n \to \infty$ are studied to find solutions of the so-called Peacock problem [63] where the marginals for a continuous-time martingale are prescribed; see also [60] and [75] for other continuous-time results with full marginal constraint. Early contributions related to the continuous-time martingale transport problem include [47, 48, 52, 92, 102, 106].

The remainder of the paper is organized as follows. Section 1.2 fixes basic terminology and recalls the necessary results from the one-step case. In Section 1.3, we characterize the polar structure of $\mathcal{M}(\boldsymbol{\mu})$. Section 1.4 introduces and analyzes the space that is the domain of the dual problem in Section 1.5, where we state the duality theorem and the monotonicity principle. Section 1.6 introduces left-monotone transports by the shadow construction and Section 1.7 develops the equivalent characterizations in terms of support and optimality properties. The (non-)uniqueness of left-monotone transports is discussed in Section 1.8. We conclude with the analysis of the problem with unconstrained intermediate marginals in Section 1.9.

1.2 Preliminaries

Throughout this paper, μ_t, μ, ν denote finite measures on $\mathbb{R}$ with finite first moment, the total mass not necessarily being normalized. Generalizing the notation from the Introduction to a vector $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ of such measures, we will write $\Pi(\boldsymbol{\mu})$ for the set of couplings; that is, measures P on $\mathbb{R}^{n+1}$ such that $P \circ X_t^{-1} = \mu_t$ for $0 \leq t \leq n$ where $X = (X_0, \ldots, X_n) : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is the identity. Moreover, $\mathcal{M}(\boldsymbol{\mu})$ is the subset of all $P \in \Pi(\boldsymbol{\mu})$ that are martingales, meaning that

$$\int X_s \mathbf{1}_A(X_0, \ldots, X_s) dP = \int X_t \mathbf{1}_A(X_0, \ldots, X_s) dP$$

for all $s \leq t$ and Borel sets $A \in \mathfrak{B}(\mathbb{R}^{s+1})$.

We denote by $\mathbb{F} = \{\mathfrak{F}_t\}_{0 \leq t \leq n}$ the canonical filtration $\mathfrak{F}_t := \sigma(X_0, \ldots, X_t)$. As usual, an $\mathbb{F}$-predictable process $H = \{H_t\}_{1 \leq t \leq n}$ is a sequence of real functions on $\mathbb{R}^{n+1}$ such that H_t is $\mathfrak{F}_{t-1}$-measurable; i.e. $H_t = h_t(X_0, \ldots, X_{t-1})$ for some Borel-measurable $h_t : \mathbb{R}^t \to \mathbb{R}$.

Given an $\mathbb{F}$-predictable process H, the discrete stochastic integral $\{(H \cdot X)_t\}_{0 \le t \le n}$ is defined by

$$(H \cdot X)_t := \sum_{s=1}^{t} H_s \cdot (X_s - X_{s-1}).$$

If X is a martingale under some measure P, then $H \cdot X$ is a generalized (not necessarily integrable) martingale in the sense of generalized conditional expectations; cf. [70, Proposition 1.64].

We say that $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ is in convex order if $\mu_{t-1} \le_c \mu_t$ for all $1 \le t \le n$; that is, $\mu_{t-1}(\phi) \le \mu_t(\phi)$ for any convex function $\phi : \mathbb{R} \to \mathbb{R}$. This implies that μ_{t-1} and μ_t have the same total mass. The order can also be characterized by the potential functions

$$u_{\mu_t} : \mathbb{R} \to \mathbb{R}, \quad u_{\mu_t}(x) := \int |x - y| \, \mu_t(dy).$$

The following properties are elementary:

(i) u_{μ_t} is nonnegative and convex,

(ii) $\partial^+ u_{\mu_t}(x) - \partial^- u_{\mu_t}(x) = 2\mu_t(\{x\})$,

(iii) $\lim_{|x|\to\infty} u_{\mu_t}(x) = \infty \mathbf{1}_{\mu_t \neq 0}$,

(iv) $\lim_{|x|\to\infty} u_{\mu_t}(x) - \mu_t(\mathbb{R})|x - \mathrm{bary}(\mu_t)| = 0$,

where ∂^+ and ∂^- denote the right and left derivatives, respectively, and $\mathrm{bary}(\mu_t) = (\int x d\mu_t)/\mu_t(\mathbb{R})$ is the barycenter. We can therefore extend u_{μ_t} continuously to $\bar{\mathbb{R}} = [-\infty, \infty]$. The following result of Strassen is classical (cf. [104]; the last statement is obtained as e.g. in [50, Corollary 2.95]).

Proposition 1.2.1. *Let $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ be finite measures on $\mathbb{R}$ with finite first moments and equal total mass. The following are equivalent:*

(i) $\mu_0 \le_c \cdots \le_c \mu_n$,

(ii) $u_{\mu_0} \le \cdots \le u_{\mu_n}$,

(iii) $\mathcal{M}(\boldsymbol{\mu}) \neq \emptyset$,

(iv) there exist stochastic kernels $\kappa_t(x_0, \dots, x_{t-1}, dx_t)$ such that

$$\int |x_t| \, \kappa_t(x_0, \dots, x_{t-1}, dx_t) < \infty \text{ and } \int x_t \, \kappa_t(x_0, \dots, x_{t-1}, dx_t) = x_{t-1}$$

for all $(x_0, \dots, x_t) \in \mathbb{R}^t$ and $1 \le t \le n$, and

$$\mu_t = (\mu_0 \otimes \kappa_1 \otimes \cdots \otimes \kappa_n) \circ (X_t)^{-1} \quad \text{for all } 0 \le t \le n.$$

All kernels will be stochastic (i.e. normalized) in what follows. A kernel κ_t with the first property in (iv) is called *martingale kernel*.

1.2.1 The One-Step Case

For the convenience of the reader, we summarize some results from [16] and [18] for the one-step problem ($n = 1$) which will be used later on. In this section we write (μ, ν) instead of (μ_0, μ_1) for the given marginals in convex order.

Definition 1.2.2. The pair $\mu \le_c \nu$ is *irreducible* if the set $I = \{u_\mu < u_\nu\}$ is connected and $\mu(I) = \mu(\mathbb{R})$. In this situation, let J be the union of I and any endpoints of I that are atoms of ν; then (I, J) is the *domain* of $\mathcal{M}(\mu, \nu)$.

The first result is a decomposition of the transport problem into irreducible parts; cf. [16, Theorem 8.4].

Proposition 1.2.3. *Let $\mu \le_c \nu$ and let $(I_k)_{1 \le k \le N}$ be the (open) components of $\{u_\mu < u_\nu\}$, where $N \in \{0, 1, \dots, \infty\}$. Set $I_0 = \mathbb{R} \backslash \cup_{k \ge 1} I_k$ and $\mu_k = \mu|_{I_k}$ for $k \ge 0$, so that $\mu = \sum_{k \ge 0} \mu_k$. Then, there exists a unique decomposition $\nu = \sum_{k \ge 0} \nu_k$ such that*

$$\mu_0 = \nu_0 \quad \text{and} \quad \mu_k \le_c \nu_k \quad \text{for all} \quad k \ge 1,$$

and this decomposition satisfies $I_k = \{u_{\mu_k} < u_{\nu_k}\}$ for all $k \ge 1$. Moreover, any $P \in \mathcal{M}(\mu, \nu)$ admits a unique decomposition $P = \sum_{k \ge 0} P_k$ such that $P_k \in \mathcal{M}(\mu_k, \nu_k)$ for all $k \ge 0$.

We observe that the measure P_0 in Proposition 1.2.3 transports μ_0 to itself and is concentrated on $\Delta_0 := \Delta \cap I_0^2$ where $\Delta = \{(x, x) : x \in \mathbb{R}\}$ is the diagonal. Thus, the transport problem with index $k = 0$ is not actually an irreducible one, but we shall nevertheless refer to (I_0, I_0) as the domain of this problem. When we want to emphasize the distinction, we call (I_0, I_0) the *diagonal domain* and $(I_k, J_k)_{k\geq 1}$ the *irreducible domains* of $\mathcal{M}(\mu,\nu)$. Similarly, the sets $V_k := I_k \times J_k$, $k \geq 1$ will be called the *irreducible components* and $V_0 := \Delta_0$ will be called the *diagonal component* of $\mathcal{M}(\mu,\nu)$. This terminology refers to the following result of [18, Theorem 3.2] which essentially states that the components are the only sets that can be charged by a martingale transport. We call a set $B \subseteq \mathbb{R}^2$ $\mathcal{M}(\mu,\nu)$-polar if it is P-null for all $P \in \mathcal{M}(\mu,\nu)$, where a nullset is, as usual, any set contained in a Borel set of zero measure.

Proposition 1.2.4. *Let $\mu \leq_c \nu$ and let $B \subseteq \mathbb{R}^2$ be a Borel set. Then B is $\mathcal{M}(\mu,\nu)$-polar if and only if there exist a μ-nullset N_μ and a ν-nullset N_ν such that*

$$B \subseteq (N_\mu \times \mathbb{R}) \cup (\mathbb{R} \times N_\nu) \cup \left(\bigcup_{k\geq 0} V_k \right)^c .$$

The following result of [18, Lemma 3.3] will also be useful; it is the main ingredient in the proof of the preceding proposition.

Lemma 1.2.5. *Let $\mu \leq_c \nu$ be irreducible and let π be a finite measure on $\mathbb{R}^2$ whose marginals π_1, π_2 satisfy*[3] *$\pi_1 \leq \mu$ and $\pi_2 \leq \nu$. Then, there exists $P \in \mathcal{M}(\mu,\nu)$ such that P dominates π in the sense of absolute continuity.*

1.3 The Polar Structure

The goal of this section is to identify all obstructions to martingale transports imposed by the marginals $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$, and thus, conversely, the sets that can indeed be charged. We recall that a subset B of $\mathbb{R}^{n+1}$ is called $\mathcal{M}(\boldsymbol{\mu})$-polar if it is a P-nullset for all $P \in \mathcal{M}(\boldsymbol{\mu})$. The result for the one-step case in Proposition 1.2.4 already exhibits an obvious type of polar set $B \subseteq \mathbb{R}^{n+1}$: if for some t there is an $\mathcal{M}(\mu_{t-1}, \mu_t)$-polar set $B' \subseteq \mathbb{R}^2$ such that $B \subseteq \mathbb{R}^{t-1} \times B' \times \mathbb{R}^{n-t}$, then B must be $\mathcal{M}(\boldsymbol{\mu})$-polar. The following shows that unions of such sets are in fact the only polar sets of $\mathcal{M}(\boldsymbol{\mu})$.

[3]By $\pi_1 \leq \mu$ we mean that $\pi_1(A) \leq \mu(A)$ for every Borel set $A \subseteq \mathbb{R}$.

Theorem 1.3.1 (Polar Structure)**.** *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order. Then a Borel set $B \subseteq \mathbb{R}^{n+1}$ is $\mathcal{M}(\boldsymbol{\mu})$-polar if and only if there exist μ_t-nullsets N_t such that*

$$B \;\subseteq\; \bigcup_{t=0}^{n}(X_t)^{-1}(N_t) \;\cup\; \bigcup_{t=1}^{n}(X_{t-1}, X_t)^{-1}\left(\bigcup_{k\geq 0} V_k^t\right)^c \tag{1.3.1}$$

where $(V_k^t)_{k\geq 1}$ are the irreducible components of $\mathcal{M}(\mu_{t-1}, \mu_t)$ and V_0^t is the corresponding diagonal component.

Before stating the proof, we introduce some additional terminology. The second part of (1.3.1) can be expressed as

$$\begin{aligned}\bigcup_{t=1}^{n}(X_{t-1}, X_t)^{-1}\left(\bigcup_{k\geq 0} V_k^t\right)^c &= \left(\bigcap_{t=1}^{n}\bigcup_{k\geq 0}(X_{t-1}, X_t)^{-1}(V_k^t)\right)^c \\ &= \left(\bigcup_{k_1,\dots,k_n\geq 0}\;\bigcap_{t=1}^{n}(X_{t-1}, X_t)^{-1}(V_{k_t}^t)\right)^c.\end{aligned} \tag{1.3.2}$$

For every $\boldsymbol{k} = (k_1, \dots, k_n)$, the set

$$V_{\boldsymbol{k}} = \bigcap_{t=1}^{n}(X_{t-1}, X_t)^{-1}(V_{k_t}^t) \subseteq \mathbb{R}^{n+1}$$

as occurring in the last expression of (1.3.2) will be referred to as an *irreducible component* of $\mathcal{M}(\boldsymbol{\mu})$; these sets are disjoint since $V_k^t \cap V_{k'}^t = \emptyset$ for $k \neq k'$. Moreover, we call their union

$$\mathcal{V} = \cup_{\boldsymbol{k}} V_{\boldsymbol{k}}$$

the *effective domain* of $\mathcal{M}(\boldsymbol{\mu})$.

Roughly speaking, an irreducible component $V_{\boldsymbol{k}}$ is a chain of irreducible components from the individual steps $(t-1, t)$. In the one-step case considered in [16, 18], it was possible and useful to decompose the transport problem into its irreducible components and study those separately to a large extent; cf. Proposition 1.2.3. This is impossible in the multistep case, as illustrated by the following example.

Example 1.3.2. Consider the two-step martingale transport problem with marginals $\mu_0 =$

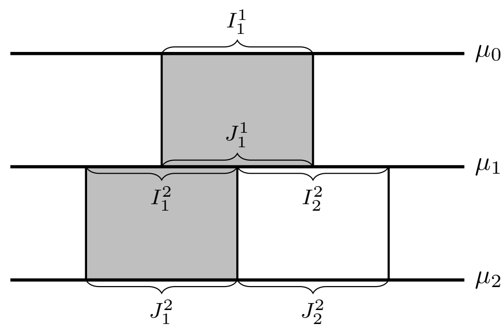


Figure 1.2: The shaded area represents $V_{\boldsymbol{k}}$ for $\boldsymbol{k} = (1, 1)$.

δ_0, $\mu_1 = \frac{1}{2}(\delta_{-1} + \delta_1)$ and $\mu_2 = \frac{1}{4}(\delta_{-2} + 2\delta_0 + \delta_2)$. Then the irreducible components are given by

$$
\begin{aligned}
V_{00} &= \{(x, x, x) : x \notin (-2, 2)\} \\
V_{01} &= \{(x, x) : x \in (-2, -1]\} \times [-2, 0] \\
V_{02} &= \{(x, x) : x \in [1, 2)\} \times [0, 2] \\
V_{10} &= (-1, 1) \times \{0\} \times \{0\} \\
V_{11} &= (-1, 1) \times [-1, 0) \times [-2, 0] \\
V_{12} &= (-1, 1) \times (0, 1] \times [0, 2].
\end{aligned}
$$

There is only one martingale transport $P \in \mathcal{M}(\boldsymbol{\mu})$, given by

$$P = \frac{1}{4}(\delta_{(0,-1,-2)} + \delta_{(0,-1,0)} + \delta_{(0,1,0)} + \delta_{(0,1,2)}).$$

While P is supported on $V_{11} \cup V_{12}$, it cannot be decomposed into two martingale parts that are supported on V_{11} and V_{12}, respectively: V_{11} and V_{12} are disjoint, but $P|_{V_{11}} = \frac{1}{4}(\delta_{(0,-1,-2)} + \delta_{(0,-1,0)})$ is not a martingale.

The main step in the proof of Theorem 1.3.1 will be the following lemma.

Lemma 1.3.3. *Let $V_{\boldsymbol{k}}$ be an irreducible component of $\mathcal{M}(\boldsymbol{\mu})$ and consider a measure π concentrated on $V_{\boldsymbol{k}}$ such that $\pi_t \leq \mu_t$ for $t = 0, \dots, n$. Then there exists a transport $P \in \mathcal{M}(\boldsymbol{\mu})$ which dominates π in the sense of absolute continuity.*

Deferring the proof, we first show how this implies the theorem.

Proof of Theorem 1.3.1. Clearly $(X_t)^{-1}(N_t)$ is $\mathcal{M}(\boldsymbol{\mu})$-polar for $t = 0, \dots, n$ and $(X_{t-1}, X_t)^{-1}\left(\bigcup_{k\geq 0} V_k^t\right)^c$ is $\mathcal{M}(\boldsymbol{\mu})$-polar for $t = 1, \dots, n$. This shows that (1.3.1) is sufficient for $B \subseteq \mathbb{R}^{n+1}$ to be $\mathcal{M}(\boldsymbol{\mu})$-polar.

Conversely, suppose that (1.3.1) does not hold; we show that B is not $\mathcal{M}(\boldsymbol{\mu})$-polar. In view of (1.3.2), by passing to a subset of B if necessary, we may assume that

$$B \subseteq \mathcal{V} = \bigcup_{\boldsymbol{k}} V_{\boldsymbol{k}} = \bigcup_{\boldsymbol{k}} \bigcap_{t=1}^{n} (X_{t-1}, X_t)^{-1}(V_{k_t}^t).$$

We may also assume that there are no μ_t-nullsets N_t such that $B \subseteq \cup_{t=0}^n (X_t)^{-1}(N_t)$. By a result of classical optimal transport [12, Proposition 2.1], this entails that B is not $\Pi(\boldsymbol{\mu})$-polar; i.e. we can find a measure $\rho \in \Pi(\boldsymbol{\mu})$ such that $\rho(B) > 0$.

We now write $B = \bigcup_{\boldsymbol{k}} B \cap V_{\boldsymbol{k}}$. As $\rho(B) = \sum_{\boldsymbol{k}} \rho(B \cap V_{\boldsymbol{k}}) > 0$, we can find some $\boldsymbol{k}$ such that $\rho(B \cap V_{\boldsymbol{k}}) > 0$. But then $\pi := \rho|_{V_{\boldsymbol{k}}}$ satisfies the assumptions of Lemma 1.3.3 which yields $P \in \mathcal{M}(\boldsymbol{\mu})$ such that $P \gg \pi$. In particular, $P(B) > 0$ and B is not $\mathcal{M}(\boldsymbol{\mu})$-polar. □

1.3.1 Proof of Lemma 1.3.3

The reasoning for Lemma 1.3.3 follows an induction on the number n of time steps; its rigorous formulation requires a certain amount of control over subsequent steps of the transport problem. Thus, we first state a more quantitative version of (the core part of) the lemma that is tailored to the inductive argument.

Definition 1.3.4. Let $\boldsymbol{\mu}$ be in convex order and $\mathcal{V}$ the effective domain of $\mathcal{M}(\boldsymbol{\mu})$. We say that a finite measure π has a *compact support family* if there are disjoint compact product sets[4] $K_1, \dots, K_m \subseteq \mathcal{V}$ with $\pi(\cup_i K_i) = \pi(\mathbb{R}^{n+1})$ such that $K_i \subseteq V_{\boldsymbol{k}_i}$ for some irreducible component $V_{\boldsymbol{k}_i}$ for all $i = 1, \dots, m$.

Definition 1.3.5. Let $\boldsymbol{\mu}$ be in convex order, $t \leq n$ and $\sigma \leq \mu_t$ a finite measure on $\mathbb{R}$. If $t = n$, we say that σ is *diagonally compatible* (with $\boldsymbol{\mu}$) if there is a finite family of compact sets $L_1, \dots, L_m \subseteq \mathbb{R}$ with $\sigma(\cup_i L_i) = \sigma(\mathbb{R})$. Whereas if $t < n$, we require in addition that for every i, either (a) $L_i \subseteq I_k$ for some irreducible component (I_k, J_k) of $\mathcal{M}(\mu_t, \mu_{t+1})$ or

[4]By a compact product set we mean a set $K = A_0 \times \cdots \times A_n$ where each $A_t \subseteq \mathbb{R}$ is compact.

(b) $L_i \subseteq I_0$ and there is $t+1 \leq t' \leq n$ such that $L_i \subseteq I_0^s$ for the diagonal components of $\mathcal{M}(\mu_s, \mu_{s+1})$ for all $t \leq s < t'$ and $L_i \subseteq I_k^{t'}$ for some (non-diagonal) irreducible component $(I_k^{t'}, J_k^{t'})$ of $\mathcal{M}(\mu_{t'}, \mu_{t'+1})$, where we set $I_k^n = J_k^n = \mathbb{R}$ for notational convenience.

Lemma 1.3.6. *Let $t < n$ and let $L \subseteq I_0$ be a compact interval contained in the diagonal component of $\mathcal{M}(\mu_t, \mu_{t+1})$ such that $\mu_t(L) > 0$. There exist a compact interval $L' \subseteq L$ with $\mu_t(L') > 0$ and $t+1 \leq t' \leq n$ such that $L' \subseteq I_0^s$ for the diagonal component of $\mathcal{M}(\mu_s, \mu_{s+1})$ for all $t \leq s < t'$ and $L' \subseteq I_k^{t'}$ for some (non-diagonal) irreducible component $(I_k^{t'}, J_k^{t'})$ of $\mathcal{M}(\mu_{t'}, \mu_{t'+1})$, where we again set $I_k^n = J_k^n = \mathbb{R}$ for notational convenience.*

Proof. The statement is trivially satisfied for $t = n-1$ as we can just take $L' = L$. For $t < n-1$, consider the family of irreducible components (I_k^{t+1}, J_k^{t+1}) of $\mathcal{M}(\mu_{t+1}, \mu_{t+2})$. We distinguish three cases.

(i) First, consider the case where $L \cap I_k^{t+1} = \emptyset$ for all $k \geq 1$, then L is contained in the diagonal component of $\mathcal{M}(\mu_{t+1}, \mu_{t+2})$.

(ia) If $L = \{x\}$ consists of a single point with positive mass, then we can conclude by induction from the result for $t+1$.

(ib) If no endpoint of L is on the boundary of some component I_k^t, then observe that $\mu_t|_L = \mu_{t+1}|_L$. We can find $L' \subseteq L$ from the statement of the lemma for $t+1$. Then L' gives the result as $\mu_t(L') = \mu_{t+1}(L') > 0$.

(ic) If L contains more than one point, and also the endpoint of some component I_k^t. When this endpoint x has positive point mass, we can set $L' = \{x\}$ and conclude as in (ia). If the endpoint has zero mass, we can find $\bar{L} \subseteq L$ compact with $\mu_t(\bar{L}) > 0$ that does not contain this endpoint and argue as in (ib). (Observe that there might be at most two endpoints.)

(ii) Next, let $k \geq 1$ be such that $\mu_{t+1}(L \cap I_k^{t+1}) > 0$ (and in particular $L \cap I_k^{t+1} \neq \emptyset$). Then we can find a compact interval $L' \subseteq L \cap I_k^{t+1}$ such that $\mu_t(L') > 0$ and we directly see that L' satisfies the statement of the lemma.

(iii) Finally, suppose that there is $k \geq 1$ with $L \cap I_k^{t+1} \neq \emptyset$ but $\mu_t(L \cap I_k^{t+1}) = 0$. In particular this means that $L \not\subseteq I_k^{t+1}$. It furthermore means that $I_k^{t+1} \not\subseteq L$, as otherwise $\mu_{t+1}(I_k^{t+1}) = \mu_t(I_k^{t+1}) = \mu_t(L \cap I_k^{t+1}) = 0$ which contradicts the definition of I_k^{t+1}. As L is a compact interval and I_k^{t+1} is an open interval, we have that $L \backslash I_k^{t+1}$ is a compact interval and $\mu_t(L \backslash I_k^{t+1}) = \mu_t(L) > 0$. Notice that there can be at most two such components I_k^{t+1} for fixed L and we will be in case (i) after removing both of them if necessary. □

Lemma 1.3.7. *Let $t \leq n$ and let $J \subseteq \mathbb{R}$ be an interval such that $\mu_t(J) > 0$. Then we can find a compact interval $K \subseteq J$ with $\mu_t(K) > 0$ such that $\mu_t|_K$ is diagonally compatible.*

Proof. The case $t = n$ is trivial. Thus, let $t < n$. We consider the family $\{I_k\}_{k \geq 1}$ of open sets corresponding to the irreducible components of $\mathcal{M}(\mu_t, \mu_{t+1})$ and distinguish two cases.

(i) There is some $k \geq 1$ such that $\mu_t(I_k \cap J) > 0$. In this case, we can choose a compact interval $K \subseteq I_k \cap J$ such that $\mu_t(K) > 0$.

(ii) Now suppose that $\mu_t(I_k \cap J) = 0$ for all $k \geq 1$. Then we first notice that there are at most two components I_{k_1}, I_{k_2} so that $I_{k_i} \cap J \neq \emptyset$ and $J \backslash (I_{k_1} \cup I_{k_2})$ is still a nonempty interval with positive μ_t-mass, since I_k cannot be contained in J. We can therefore assume without loss of generality that $J \subseteq I_0$ and is compact. Now we can apply Lemma 1.3.6 to find a subinterval $K \subseteq J$ such that $\mu_t|_K$ is diagonally compatible. □

Lemma 1.3.8. *Let $t \leq n$ and let π be a measure on $\mathbb{R}^{t+1}$ that has a compact support family with respect to $\mu_0, \ldots, \mu_t$ and satisfies $\pi_s \leq \mu_s$ for $s \leq t$. In addition, suppose that π_t is diagonally compatible.*

Then there is a martingale measure Q on $\mathbb{R}^{t+1}$ that dominates π in the sense of absolute continuity and has a compact support family with respect to $\mu_0, \ldots, \mu_t$ and satisfies $Q_s \leq \mu_s$ for $s \leq t$. In addition, Q_t can be chosen to be diagonally compatible. Finally, Q can be chosen such that $dQ = g d\pi + d\sigma$ where the density g is bounded and the measure σ is singular with respect to π.

Proof. We proceed by induction on t. For $t = 0$ there is nothing to prove; we can set $Q = \pi$.

Consider $t \geq 1$ and assume that the lemma has already been shown for $(t-1)$-step measures. We disintegrate

$$\pi = \pi' \otimes \kappa(x_0, \ldots, x_{t-1}, dx_t) \tag{1.3.3}$$

and observe that π' satisfies the conditions of the lemma. In particular, π'_{t-1} must be diagonally compatible: the compact sets that it is supported on are either contained in irreducible components of $\mathcal{M}(\mu_{t-1}, \mu_t)$ or in the diagonal component. Any such compact subset of the diagonal component of $\mathcal{M}(\mu_{t-1}, \mu_t)$ must correspond to one of the finitely many compact sets in the support of π_t so that they inherit the compatibility property from these sets.

By the induction assumption, we then find a martingale measure $Q' \gg \pi'$ on $\mathbb{R}^t$ with the stated properties. In particular, the marginal Q'_{t-1} is diagonally compatible with $\boldsymbol{\mu}$.

Again, let $\{I_k\}_{k\geq 1}$ be the open intervals from the irreducible domains (I_k, J_k) of $\mathcal{M}(\mu_{t-1}, \mu_t)$ and let I_0 denote the corresponding diagonal domain. We shall construct a martingale kernel $\hat{\kappa}$ by suitably manipulating κ. Let us observe that since π is concentrated on $\mathcal{V}$ and has a compact support family with respect to $\mu_0, \dots, \mu_t$, the following hold for π'-a.e. $\boldsymbol{x} = (x_0, \dots, x_{t-1}) \in \mathbb{R}^t$ and a finite family of compact sets L_i with properties (a) or (b) from Definition 1.3.5:

- $\kappa(\boldsymbol{x}, \cdot) = \delta_{x_{t-1}}$ whenever $x_{t-1} \in I_0$,
- $\kappa(\boldsymbol{x}, \cdot)$ is concentrated on some L_i with $L_i \subseteq J_k$ for $x_{t-1} \in I_k$ with $k \geq 1$ and $Q'_{t-1}(I_k) > 0$.

By changing κ on a π'-nullset, we may assume that these two properties hold for all $\boldsymbol{x} \in \mathbb{R}^t$.

Step 1. Next, we argue that we may change Q' and κ such that the marginal $(Q' \otimes \kappa)_t = (Q' \otimes \kappa) \circ X_t^{-1}$ satisfies

$$(Q' \otimes \kappa)_t \leq \mu_t. \tag{1.3.4}$$

Indeed, recall that $dQ' = dQ'_{abs} + d\sigma' = g' d\pi' + d\sigma'$ where the density g' is bounded and σ' is singular with respect to π'. Using the Lebesgue decomposition theorem, we find a Borel set $A \subseteq \mathbb{R}^t$ such that $\sigma'(A) = \sigma'(\mathbb{R}^t)$ and $\pi'(A) = 0$. By scaling Q' with a constant we may assume that $g' \leq 1/2$. As $\pi_t \leq \mu_t$, the marginal $(Q'_{abs} \otimes \kappa)_t$ is then bounded by $\frac{1}{2}\mu_t$, and it remains to bound $(\sigma' \otimes \kappa)_t$ in the same way.

Note that $Q'_{t-1} \leq \mu_{t-1}$ implies $\sigma'_{t-1} \leq \mu_{t-1}$. We may change κ arbitrarily on the set A without invalidating (1.3.3). Indeed, for each irreducible component (I_k, J_k) of $\mathcal{M}(\mu_{t-1}, \mu_t)$ we choose and fix a compact interval $K_k \subseteq J_k$ with $\mu_t(K_k) > 0$ such that $\mu_t|_{K_k}$ is diagonally compatible; this is possible by Lemma 1.3.7. For $\boldsymbol{x} = (x_0, \dots, x_{t-1}) \in A$ such that $x_{t-1} \in I_k$ we then define

$$\kappa(\boldsymbol{x}, \cdot) := \frac{1}{\mu_t(K_k)} \mu_t|_{K_k}.$$

Set $\epsilon_k = \mu_t(K_k)/\mu_{t-1}(I_k)$. Then

$$\epsilon := \inf_{k:\, Q'_{t-1}(I_k)>0} \epsilon_k \wedge 1$$

is strictly positive because there are only finitely many k with $Q'_{t-1}(I_k) > 0$ (this is the purpose of the induction assumption that Q'_{t-1} is diagonally compatible). As $\sigma'_{t-1} \le \mu_{t-1}$, we may scale Q' once again to obtain $\sigma'_{t-1} \le \frac{\epsilon}{6}\mu_{t-1}$. We now have

$$(\sigma'|_{\mathbb{R}^{t-1}\times I_k} \otimes \kappa)_t = \sigma'_{t-1}(I_k)\frac{1}{\mu_t(K_k)}\mu_t|_{K_k} \le \frac{\epsilon}{6}\frac{\mu_{t-1}(I_k)}{\mu_t(K_k)}\mu_t|_{K_k} \le \frac{1}{6}\mu_t|_{K_k}.$$

For the diagonal domain I_0 the corresponding inequality holds because we have $\kappa(\boldsymbol{x},\cdot) = \delta_{x_{t-1}}$ for $x_{t-1} \in I_0$ and $\sigma'_{t-1}|_{I_0} \le \frac{1}{6}\mu_{t-1}|_{I_0} \le \frac{1}{6}\mu_t|_{I_0}$. As a consequence, we have $(\sigma' \otimes \kappa)_t \le \frac{1}{2}\mu_t$ as desired, so that we may assume (1.3.4) in what follows.

Step 2. We now construct a martingale kernel $\hat\kappa$ such that $Q = Q' \otimes \hat\kappa$ has the required properties. For a fixed irreducible component (I_k, J_k) we have that $Q'_{t-1}|_{I_k} = Q'_{t-1}|_K$ for some compact $K \subseteq I_k$. We can find compact intervals $B^-, B^+ \subseteq J_k$ with $\mu_t(B^-) > 0$ and $\mu_t(B^+) > 0$ such that B^- is to the left of K and B^+ is to the right of K, in the sense that $x < y < z$ for $x \in B^-$, $y \in K$ and $z \in B^+$. By Lemma 1.3.7, we can further assume that we have $B^+ \subseteq I^t_k$ and $B^- \subseteq I^t_{k'}$ for some $k, k' \ge 0$, where $(I^t_l)_{l\ge 0}$ belong to the components of $\mathcal{M}(\mu_t, \mu_{t+1})$, and that $\mu_t|_{B^\pm}$ is diagonally compatible.

Next, we define two nonnegative functions $\boldsymbol{x} \mapsto \varepsilon^-(\boldsymbol{x}), \varepsilon^+(\boldsymbol{x})$ for $\boldsymbol{x} = (x_0, \dots, x_{t-1}) \in \mathbb{R}^{t-1} \times K$ as follows:

- for $\boldsymbol{x}$ such that $\mathrm{bary}(\kappa(\boldsymbol{x},\cdot)) < x_{t-1}$, let ε^+ be the unique number such that $\kappa(\boldsymbol{x},\cdot) + \varepsilon^+(\boldsymbol{x}) \cdot \mu_t|_{B^+}$ has barycenter x_{t-1},
- for $\boldsymbol{x}$ such that $\mathrm{bary}(\kappa(\boldsymbol{x},\cdot)) > x_{t-1}$, let ε^- be the unique number such that $\kappa(\boldsymbol{x},\cdot) + \varepsilon^-(\boldsymbol{x}) \cdot \mu_t|_{B^-}$ has barycenter x_{t-1},
- $\varepsilon^\pm(\boldsymbol{x}) = 0$ otherwise.

Observe that these numbers always exist because B^- and B^+ have positive mass and positive distance from the points $x_{t-1} \in K$. We now define the martingale kernel $\hat{\kappa}$ by

$$\hat{\kappa}(\boldsymbol{x}) := c(\varepsilon^- \cdot \mu_t|_{B^-} + \kappa + \varepsilon^+ \cdot \mu_t|_{B^+})$$

where $0 < c \leq 1$ is a normalizing constant such that $\hat{\kappa}$ is again a stochastic kernel. We also define $\hat{\kappa}(\boldsymbol{x}) = \kappa(\boldsymbol{x})$ for $\boldsymbol{x}$ on the diagonal domain.

For each $k \geq 1$, let $B_k^{\pm}$ denote the sets associated with I_k as above. Once again, the number

$$C := \frac{1}{3} \inf_{k:\, Q'_{t-1}(I_k)>0} [\mu_t(B_k^-) \wedge \mu_t(B_k^+)]$$

is strictly positive because there are only finitely many k with $Q'_{t-1}(I_k) > 0$. We can now define

$$Q := C \cdot (Q' \otimes \hat{\kappa}).$$

Then Q is a martingale transport whose marginals satisfy $Q_s \leq Q'_s \leq \mu_s$ for $0 \leq s \leq t-1$ whereas $Q_t \leq \mu_t$ by (1.3.4), the construction of $\hat{\kappa}$ and the choice of C; indeed, for every $x_{t-1} \in I_k^t$ we have

$$\begin{aligned} 3C\hat{\kappa}(\boldsymbol{x}) &\leq 3C\varepsilon^- \cdot \mu_t|_{B^-} + 3C\kappa + 3C\varepsilon^+ \cdot \mu_t|_{B^+} \\ &\leq \mu_t|_{B^-} + \kappa + \mu_t|_{B^+} \leq 2\mu_t + \kappa. \end{aligned}$$

To see that Q_t is diagonally compatible, observe that Q_t is supported by a finite family of compact sets consisting of the following:

- a finite family of compact sets $\bar{L}_i \subseteq I_0$ such that $Q'_{t-1}|_{\bar{L}_i}$ is diagonally compatible (from the induction hypothesis that Q'_{t-1} is diagonally compatible),
- a finite family of compact sets $L_i \subseteq J_k$ for some $k \geq 1$ with $Q'_{t-1}(I_k) > 0$ such that $Q_t|_{L_i} \leq \mu_t|_{L_i}$ is diagonally compatible, and
- the sets $B_k^{\pm}$ for the finitely many k such that $Q'_{t-1}(I_k) > 0$, where $Q_t|_{B_k^{\pm}} \leq \mu_t|_{B_k^{\pm}}$ is diagonally compatible.

It remains to check that Q has the required decomposition with respect to π. Indeed, $\hat{\kappa}$

can be decomposed as

$$\hat{\kappa} = c\kappa + (1-c)\kappa^{\perp}$$

where $\kappa^{\perp}$ is singular to κ. Recalling the decomposition $Q' = Q'_{abs} + \sigma'$, we then have

$$Q' \otimes \hat{\kappa} = cQ'_{abs} \otimes \kappa \ + \ (1-c)Q'_{abs} \otimes \kappa^{\perp} \ + \ \sigma' \otimes \hat{\kappa}.$$

The last two terms are singular with respect to $\pi = \pi' \otimes \kappa$, and the first term is absolutely continuous with bounded density. $\square$

Proof of Lemma 1.3.3. Let π be a measure with marginals $\pi_t \leq \mu_t$ for all t which is concentrated on some irreducible component $V = V_{\boldsymbol{k}}$ and thus, in particular, on the effective domain $\mathcal{V}$.

Step 1. We first decompose $\pi = \sum_{m=1}^{\infty} \pi^m$ such that each π^m satisfies the requirements of Lemma 1.3.8 with $t = n$.

Indeed, let $V = \cap_{t=1}^{n} (X_{t-1}, X_t)^{-1}(V^t_{k_t})$ and suppose first that $k_t \neq 0$ for $1 \leq t \leq n$. Then, we can write V as a product of nonempty intervals: $V = A_0 \times \cdots \times A_n$ where $A_0 = I^1_{k_1}$, $A_n = J^n_{k_n}$ and $A_t = J^t_{k_t} \cap I^{t+1}_{k_{t+1}}$ for $1 < t < n$. Thus, we can choose increasing families of compact intervals K^m_t such that $A_t = \cup_{m \geq 1} K^m_t$ for all t. Setting $\pi^1 := \pi|_{\prod_{t=0}^{n} K^1_t}$ and $\pi^m := \pi|_{\prod_{t=0}^{n} K^m_t \setminus \prod_{t=0}^{n} K^{m-1}_t}$ for $m > 1$ yields the required decomposition.

If $k_t = 0$ for one or more $1 \leq t \leq n$, we have $V \subseteq A_0 \times \cdots \times A_n$, where A_t is defined as above when $k_t \neq 0 \neq k_{t+1}$ but we use $\mathbb{R}$ instead of $J^t_{k_t}$ when $k_t = 0$ and $\mathbb{R}$ instead of $I^{t+1}_{k_{t+1}}$ when $k_{t+1} = 0$. After these modifications, π^m can be defined as above; recall that diagonal components are always closed.

Step 2. For each of the measures π^m, Lemma 1.3.8 yields a martingale measure $Q^m \gg \pi^m$ with the properties stated in the lemma. In particular, each Q^m has a compact support family. We show below that there exist $P^m \in \mathcal{M}(\boldsymbol{\mu})$ such that $P^m \gg Q^m$, and then $P := \sum 2^{-m} P^m$ satisfies $P \in \mathcal{M}(\boldsymbol{\mu})$ and $P \gg \pi$ as desired.

To complete the proof, it remains to show that for fixed $m \geq 1$ there exist $0 < \epsilon < 1$ and $\bar{Q}^m \in \mathcal{M}(\boldsymbol{\mu} - \epsilon(Q^m_0, \ldots, Q^m_n))$, as we may then conclude by setting $P^m := \epsilon Q^m + \bar{Q}^m \in \mathcal{M}(\boldsymbol{\mu})$. By Proposition 1.2.1, the set $\mathcal{M}(\boldsymbol{\mu} - \epsilon(Q^m_0, \ldots, Q^m_n))$ is nonempty if the marginals are in convex order, or equivalently if the potential functions satisfy

$$u_{\mu_{t-1}} - \epsilon u_{Q^m_{t-1}} \leq u_{\mu_t} - \epsilon u_{Q^m_t} \tag{1.3.5}$$

for $t = 1, \dots, n$. Thus, it suffices to find $\epsilon > 0$ with this property for fixed t, and we have reduced to a question about a one-step martingale transport problem. Indeed, we have $u_{\mu_{t-1}} \leq u_{\mu_t}$ on $\mathbb{R}$. Since Q^m has a compact support family and in particular is supported by $\mathcal{V}$, there is a finite collection of compact sets $K_j \subseteq \mathbb{R}$ such that each K_j is contained in one of the intervals $I^{t-1}_{k_j}$ from the decomposition of (μ_{t-1}, μ_t) into irreducible components, Q^m transports mass from K_j to itself for each j, and Q^m is the identical Monge transport on the complement $(\cup_j K_j)^c$. On each K_j, Steps (a) and (b) in the proof of [18, Lemma 3.3] yield $\epsilon > 0$ such that (1.3.5) holds on K_j, and we can choose $\epsilon > 0$ independently of j since there are finitely many j. On the other hand, (1.3.5) trivially holds on $(\cup_j K_j)^c$ since $u_{Q^m_{t-1}} = u_{Q^m_t}$ on that set. This completes the proof. □

1.4 The Dual Space

In this section we introduce the domain of the dual optimization problem and show that it has a certain closedness property. The latter will be crucial for the duality theorem in the subsequent section.

We shall need a generalized notion of integrability for the elements of the dual space. To this end, we first recall the integral for concave functions as detailed in [18, Section 4.1].

Definition 1.4.1. Let $\mu \leq_c \nu$ be irreducible with domain (I, J) and let $\chi : J \to \mathbb{R}$ be a concave function. We define

$$(\mu - \nu)(\chi) := \frac{1}{2} \int_I (u_\mu - u_\nu) d\chi'' + \int_{J \setminus I} |\Delta \chi| d\nu \in [0, \infty]$$

where $-\chi''$ is the (locally finite) second derivative measure of $-\chi$ on I and $|\Delta \chi|$ is the absolute magnitude of the jumps of χ at the boundary points $J \setminus I$.

Remark 1.4.2. As shown in [18, Lemma 4.1], this integral is well-defined and satisfies

$$(\mu-\nu)(\chi)=\int_I \left[\chi(x)-\int_J \chi(y)\,\kappa(x,dy)\right]\mu(dx)$$

for any $P=\mu\otimes\kappa\in\mathcal{M}(\mu,\nu)$. Moreover, it coincides with the difference $\mu(\chi)-\nu(\chi)$ of the usual integrals when $\chi\in L^1(\mu)\cap L^1(\nu)$.

For later reference, we record two more properties of the integral.

Lemma 1.4.3. *Let $\mu\leq_c\nu$ be irreducible with domain (I,J) and let $\chi:J\to\mathbb{R}$ be concave.*

(i) Assume that I has a finite right endpoint r and $\chi(a)=\chi'(a)=0$ for some $a\in I$. Then $\chi\leq 0$ and $\chi\mathbf{1}_{[a,\infty)}$ is concave. If ν has an atom at r, then

$$\chi(r)\geq-\frac{C}{\nu(\{r\})}(\mu-\nu)(\chi\mathbf{1}_{[a,\infty)})$$

for a constant $C\geq 0$ depending only on μ,ν.

(ii) For $a,b\in\mathbb{R}$, the concave function $\bar{\chi}(x):=\chi(x)+ax+b$ satisfies

$$(\mu-\nu)(\bar{\chi})=(\mu-\nu)(\chi).$$

Proof. The first part is [18, Remark 4.6] and the second part follows directly from $\bar{\chi}''=\chi''$ and $\Delta\bar{\chi}=\Delta\chi$. ☐

Let us now return to the multistep case with a vector $\boldsymbol{\mu}=(\mu_0,\dots,\mu_n)$ of measures in convex order and introduce $\boldsymbol{\mu}(\boldsymbol{\phi}):=\sum_{t=0}^n\mu_t(\phi_t)$ in cases where we do not necessarily have $\phi_t\in L^1(\mu_t)$. As mentioned previously, in contrast to [18], the multistep transport problem does not decompose into irreducible components, forcing us to directly give a global definition of the integral.

Definition 1.4.4. Let $\boldsymbol{\phi}=(\phi_0,\dots,\phi_n)$ be a vector of Borel functions $\phi_t:\mathbb{R}\to\bar{\mathbb{R}}$. A vector $\boldsymbol{\chi}=(\chi_1,\dots,\chi_n)$ of Borel functions $\chi_t:\mathbb{R}\to\mathbb{R}$ is called a *concave moderator* for $\boldsymbol{\phi}$ if for $1\leq t\leq n$,

(i) $\chi_t|_J$ is concave for every domain (I,J) of an irreducible component of $\mathcal{M}(\mu_{t-1},\mu_t)$,

(ii) $\chi_t|_{I_0} \equiv 0$ for the diagonal domain I_0 of $\mathcal{M}(\mu_{t-1}, \mu_t)$,

(iii) $\phi_t - \chi_{t+1} + \chi_t \in L^1(\mu_t)$,

where $\chi_{n+1} \equiv 0$. We also convene that $\chi_0 \equiv 0$. The *moderated integral* of $\boldsymbol{\phi}$ is then defined by

$$\boldsymbol{\mu}(\boldsymbol{\phi}) := \sum_{t=0}^{n} \mu_t(\phi_t - \chi_{t+1} + \chi_t) + \sum_{t=1}^{n} \sum_{k \geq 1} (\mu_{t-1} - \mu_t)^k (\chi_t) \in (-\infty, \infty], \tag{1.4.1}$$

where $(\mu_{t-1} - \mu_t)^k(\chi_t)$ denotes the integral of Definition 1.4.1 on the k-th irreducible component of $\mathcal{M}(\mu_{t-1}, \mu_t)$.

Remark 1.4.5. The moderated integral is independent of the choice of the moderator $\boldsymbol{\chi}$. To see this, consider a second moderator $\tilde{\boldsymbol{\chi}}$ for $\boldsymbol{\phi}$; then we have $(\tilde{\chi}_{t+1} - \chi_{t+1}) - (\tilde{\chi}_t - \chi_t) \in L^1(\mu_t)$. We may assume that (1.4.1) is finite for at least one of the moderators. Using Remark 1.4.2 with arbitrary κ_t such that $\mu_{t-1} \otimes \kappa_t \in \mathcal{M}(\mu_{t-1}, \mu_t)$ for $1 \leq t \leq n$, as well as Fubini's theorem for kernels,

$$\begin{aligned}
&\sum_{t=1}^{n} \sum_{k \geq 1} (\mu_{t-1} - \mu_t)^k (\chi_t) - (\mu_{t-1} - \mu_t)^k (\tilde{\chi}_t) \\
&\quad = \int \cdots \int \sum_{t=1}^{n} \chi_t(x_{t-1}) - \chi_t(x_t)\, \kappa_n(x_{n-1}, dx_n) \cdots \kappa_1(x_0, dx_1) \mu_0(dx_0) \\
&\qquad - \int \cdots \int \sum_{t=1}^{n} \tilde{\chi}_t(x_{t-1}) - \tilde{\chi}_t(x_t)\, \kappa_n(x_{n-1}, dx_n) \cdots \kappa_1(x_0, dx_1) \mu_0(dx_0) \\
&\quad = \sum_{t=0}^{n} \mu_t((\chi_{t+1} - \tilde{\chi}_{t+1}) - (\chi_t - \tilde{\chi}_t)).
\end{aligned}$$

It now follows that (1.4.1) yields the same value for both moderators.

For later reference, we also record the following property.

Remark 1.4.6. If $\boldsymbol{\chi}$ is a concave moderator, Definition 1.4.4 (ii) implies that

$$\chi_t = \sum_{k \geq 1} \chi_t|_{I_k^t} = \sum_{k \geq 1} \chi_t|_{J_k^t}$$

where (I_k^t, J_k^t) is the k-th irreducible domain of $\mathcal{M}(\mu_{t-1}, \mu_t)$.

Next, we introduce the space of functions which have a finite integral in the moderated sense.

Definition 1.4.7. We denote by $L^c(\boldsymbol{\mu})$ the space of all vectors $\boldsymbol{\phi}$ admitting a concave moderator $\boldsymbol{\chi}$ with $\sum_{t=1}^n \sum_{k\geq 1}(\mu_{t-1}-\mu_t)^k(\chi_t) < \infty$.

It follows that $\boldsymbol{\mu}(\boldsymbol{\phi})$ is finite for $\boldsymbol{\phi} \in L^c(\boldsymbol{\mu})$, and we have $\boldsymbol{\mu}(\phi) = \sum_t \mu_t(\phi_t)$ for $\boldsymbol{\phi} \in \Pi_{t=0}^n L^1(\mu_t)$. The definition is also consistent with the expectation under martingale transports, in the following sense.

Lemma 1.4.8. *Let $\boldsymbol{\phi} \in L^c(\boldsymbol{\mu})$ and let $H = (H_1, \dots, H_n)$ be $\mathbb{F}$-predictable. If*

$$\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n$$

is bounded from below on the effective domain $\mathcal{V}$ of $\mathcal{M}(\boldsymbol{\mu})$, then

$$\boldsymbol{\mu}(\boldsymbol{\phi}) = P\left[\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n\right], \quad P \in \mathcal{M}(\boldsymbol{\mu}).$$

Proof. Let $P \in \mathcal{M}(\boldsymbol{\mu})$, let $\boldsymbol{\chi}$ be a concave moderator for $\boldsymbol{\phi}$, and assume without loss of generality that 0 is the lower bound. Using Remark 1.4.6, we have that $\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n$ equals

$$\sum_{t=0}^n (\phi_t - \chi_{t+1} + \chi_t)(X_t) + \sum_{t=1}^n \sum_{k\geq 1} (\chi_t|_{I_k^t}(X_{t-1}) - \chi_t|_{J_k^t}(X_t)) + (H \cdot X)_n \geq 0.$$

By assumption, the functions $(\phi_t - \chi_{t+1} + \chi_t)(X_t)$ are P-integrable. Therefore, the negative part of the remaining expression must also be P-integrable. Writing $P_t := P \circ (X_0, \dots, X_t)^{-1}$ and using that $(\chi_t|_{J_k^t})^+$ has linear growth, we see that for any disintegration $P = P_{n-1} \otimes \kappa_n$,

$$\begin{aligned}
&\int \left[\sum_{t=1}^n \sum_{k\geq 1} (\chi_t|_{I_k^t}(X_{t-1}) - \chi_t|_{J_k^t}(X_t)) + (H \cdot X)_n\right] \kappa_n(X_0, \dots, X_{n-1}, dX_n) \\
&\quad = \sum_{t=1}^{n-1} \sum_{k\geq 1} (\chi_t|_{I_k^t}(X_{t-1}) - \chi_t|_{J_k^t}(X_t)) + (H \cdot X)_{n-1} \\
&\qquad + \sum_{k\geq 1} \int \left[\chi_n|_{I_k^n}(X_{n-1}) - \chi_n|_{J_k^n}(X_n)\right] \kappa_n(X_0, \dots, X_{n-1}, dX_n).
\end{aligned}$$

Iteratively integrating with kernels such that $P_t = P_{t-1} \otimes \kappa_t$ and observing that we can apply Fubini's theorem to $\sum_{t=1}^n \sum_{k\geq 1} (\chi_t|_{I_k^t}(X_{t-1}) - \chi_t|_{J_k^t}(X_t)) + (H \cdot X)_n$ as its negative part is P-integrable, we obtain

$$P\left[\sum_{t=1}^{n}\sum_{k\geq 1}(\chi_t|_{I_k^t}(X_{t-1})-\chi_t|_{J_k^t}(X_t))+(H\cdot X)_n\right]=\sum_{t=1}^{n}\sum_{k\geq 1}(\mu_{t-1}-\mu_t)^k(\chi_t)$$

and the result follows. □

We can now define our dual space. It will be convenient to work with nonnegative reward functions f for the moment—we shall relax this constraint later on; cf. Remark 1.5.3.

Definition 1.4.9. Let $f:\mathbb{R}^{n+1}\to[0,\infty]$. We denote by $\mathcal{D}_{\boldsymbol{\mu}}(f)$ the set of all pairs $(\boldsymbol{\phi},H)$ where $\boldsymbol{\phi}\in L^c(\boldsymbol{\mu})$ and $H=(H_1,\dots,H_n)$ is an $\mathbb{F}$-predictable process such that

$$\sum_{t=0}^{n}\phi_t(X_t)+(H\cdot X)_n\geq f \quad \text{on} \quad \mathcal{V}.$$

By Lemma 1.4.8, the expectation of the left hand side under any $P\in\mathcal{M}(\boldsymbol{\mu})$ is given by the moderated integral $\boldsymbol{\mu}(\boldsymbol{\phi})$; this will be seen as the dual cost of $(\boldsymbol{\phi},H)$ when we consider the dual problem $\inf_{(\boldsymbol{\phi},H)\in\mathcal{D}_{\boldsymbol{\mu}}(f)}\boldsymbol{\mu}(\boldsymbol{\phi})$ in Section 1.5 below.

The following closedness property is the key result about the dual space.

Proposition 1.4.10. *Let $f^m:\mathbb{R}^{n+1}\to[0,\infty]$, $m\geq 1$ be a sequence of functions such that*

$$f^m\to f \qquad \textit{pointwise}$$

and let $(\boldsymbol{\phi}^m,H^m)\in\mathcal{D}_{\boldsymbol{\mu}}(f^m)$ be such that $\sup_m\boldsymbol{\mu}(\boldsymbol{\phi}^m)<\infty$. Then there exist $(\boldsymbol{\phi},H)\in\mathcal{D}_{\boldsymbol{\mu}}(f)$ with

$$\boldsymbol{\mu}(\boldsymbol{\phi})\leq\liminf_{m\to\infty}\boldsymbol{\mu}(\boldsymbol{\phi}^m).$$

1.4.1 Proof of Proposition 1.4.10

An attempt to prove Proposition 1.4.10 directly along the lines of [18] runs into a technical issue in controlling the concave moderators. Roughly speaking, they do not allow sufficiently many normalizations; this is related to the aforementioned fact that the multistep problem cannot be decomposed into its components. We shall introduce a generalized dual space with families of functions indexed by the components, and prove a "lifted" version of Proposition 1.4.10 in this larger space. Once that is achieved, we can infer the closedness result in the original space as well. (The reader willing to admit Proposition 1.4.10 may skip this subsection without much loss of continuity.)

Definition 1.4.11. Let $\boldsymbol{\phi} = \{\phi_t^k : 0 \leq t \leq n, k \geq 0\}$ be a family of Borel functions, consisting of one function $\phi_t^k : J_k^t \to \bar{\mathbb{R}}$ for each irreducible component (I_k^t, J_k^t) of $\mathcal{M}(\mu_{t-1}, \mu_t)$ as indexed by $k \geq 1$ and $1 \leq t \leq n$, functions $\phi_t^0 : I_0^t \to \bar{\mathbb{R}}$ for the diagonal components I_0^t indexed by $1 \leq t \leq n$, and a single function $\phi_0^0 : \mathbb{R} \to \bar{\mathbb{R}}$ for $t = 0$. Similarly, let $\boldsymbol{\chi} = \{\chi_t^k : 1 \leq t \leq n, k \geq 0\}$ be a family of functions, consisting of one concave function $\chi_t^k : J_k^t \to \mathbb{R}$ for each irreducible component (I_k^t, J_k^t) and Borel functions $\chi_t^0 : I_0^t \to \mathbb{R}$ for the diagonal components. We also convene that $\chi_0^0 \equiv 0$ and define the functions[5] $\chi_t := \sum_{k\geq 0} \chi_t^k|_{I_k^t}$ for $t = 1, \ldots, n$, as well as $\chi_{n+1} \equiv 0$.

We call $\boldsymbol{\chi}$ a *concave moderator* for $\boldsymbol{\phi}$ if for all $t = 0, \ldots, n$ and $k \geq 0$,

$$\phi_t^k + \chi_t^k - \chi_{t+1} \in L^1(\mu_t^k)$$

and the sum $\sum_{k\geq 0} \mu_t^k(\phi_t^k + \chi_t^k - \chi_{t+1})$ converges in $(-\infty, \infty]$, where μ_t^k is the second marginal of the k-th irreducible component in the decomposition of $\mathcal{M}(\mu_{t-1}, \mu_t)$ as in Proposition 1.2.3 and $\mu_0^0 \equiv \mu_0$. The generalized[6] moderated integral is then defined by

$$\boldsymbol{\mu}(\boldsymbol{\phi}) := \sum_{t=0}^{n} \sum_{k\geq 0} \mu_t^k(\phi_t^k + \chi_t^k - \chi_{t+1}) + \sum_{t=1}^{n} \sum_{k\geq 1} (\mu_{t-1} - \mu_t)^k(\chi_t^k).$$

We denote by $L^{c,g}(\boldsymbol{\mu})$ the set of all families $\boldsymbol{\phi}$ which admit a concave moderator $\boldsymbol{\chi}$ such that

$$\sum_{t=0}^{n} \sum_{k\geq 0} |\mu_t^k(\phi_t^k + \chi_t^k - \chi_{t+1})| + \sum_{t=1}^{n} \sum_{k\geq 1} (\mu_{t-1} - \mu_t)^k(\chi_t^k) < \infty.$$

For $\boldsymbol{\phi} \in L^{c,g}(\boldsymbol{\mu})$, the value of $\boldsymbol{\mu}(\boldsymbol{\phi})$ is independent of the choice of the moderator $\boldsymbol{\chi}$. This is shown similarly as in Remark 1.4.5. We can now introduce the generalized dual space.

Definition 1.4.12. Let $f : \mathbb{R}^{n+1} \to [0, \infty]$. We denote by $\mathcal{D}_{\boldsymbol{\mu}}^g(f)$ the set of all pairs $(\boldsymbol{\phi}, H)$

[5]The restriction to I_k^t is important to avoid "double counting" in the sums. Note that the intervals J may overlap at their endpoints.

[6]This integral is not related to the notion of a generalized martingale.

where $\boldsymbol{\phi} \in L^{c,g}(\boldsymbol{\mu})$, $H = (H_1, \dots, H_n)$ is $\mathbb{F}$-predictable, and

$$\sum_{t=0}^{n} \phi_t^{k_t}(x_t) + (H \cdot \boldsymbol{x})_n \geq f(\boldsymbol{x})$$

for all $\boldsymbol{x} = (x_0, \dots, x_n)$ and $\boldsymbol{k} = (k_0, \dots, k_n)$ such that $(x_{t-1}, x_t) \in (I^t_{k_t}, J^t_{k_t})$ for some (irreducible or diagonal) component[7] and $t = 1, \dots, n$.

We observe that for any $\boldsymbol{x} \in \mathcal{V}$ the corresponding $\boldsymbol{k} = (k_0, \dots, k_n)$ is uniquely defined, where the index $k_0 \equiv 0$ exists purely for notational convenience.

For later reference, the following lemma elaborates on certain degrees of freedom in choosing elements of $\mathcal{D}^g_{\boldsymbol{\mu}}(f)$.

Lemma 1.4.13. *Let $(\boldsymbol{\phi}, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)$ and let $\boldsymbol{\chi}$ be a corresponding concave moderator. Let $1 \leq t \leq n$, let (I^t_k, J^t_k) be the domain of an irreducible component of $\mathcal{M}(\mu_{t-1}, \mu_t)$ and $c_1, c_2 \in \mathbb{R}$. Introduce new families $(\tilde{\boldsymbol{\phi}}, \tilde{H})$ and $\tilde{\boldsymbol{\chi}}$ by either (i) or (ii):*

(i) Define

$$\begin{aligned} \tilde{\phi}^k_t(y) &= \phi^k_t(y) - (c_1 y - c_2), \qquad \tilde{\chi}^k_t(y) = \chi^k_t(y) + (c_1 y - c_2), \\ \tilde{\phi}^{k'}_{t-1}(x) &= \phi^{k'}_{t-1}(x) + (c_1 x - c_2)|_{I^t_k}, \qquad \tilde{\chi}^{k'}_{t-1} = \chi^{k'}_{t-1}, \\ \tilde{\phi}^{k'}_s &= \phi^{k'}_s, \qquad \tilde{\chi}^{k'}_s = \chi^{k'}_s \quad \text{for } s \notin \{t-1, t\}, \\ \tilde{H}_t &= H_t + c_1|_{X^{-1}_{t-1}(I^t_k)}, \qquad \tilde{H}_s = H_s \quad \text{for } s \neq t \end{aligned}$$

where k' runs over all components of the corresponding step in the subscript.

(ii) Define

$$\tilde{\phi}^0_t = \phi^0_t + \chi^0_t - \chi^0_{t+1}|_{I^t_0}, \qquad \tilde{\chi}^0_t = 0, \quad \text{and}$$
$$\tilde{\phi}^k_t = \phi^k_t - \chi^0_{t+1}, \qquad \tilde{\chi}^k_t = \chi^k_t \quad \text{for } k \geq 1,\ t = 0, \dots, n.$$

[7] Given an irreducible component (I, J), the notation $(x, y) \in (I, J)$ means that $x \in I, y \in J$, whereas for a diagonal component (I_0, I_0) it is to be understood as $x = y \in I_0$.

Then $(\tilde{\boldsymbol{\phi}}, \tilde{H}) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)$ *and* $\tilde{\boldsymbol{\chi}}$ *is a corresponding concave moderator. Moreover, we have*

$$\sum_{t=0}^{n} \phi_t^{k_t}(x_t) + (H \cdot \boldsymbol{x})_n = \sum_{t=0}^{n} \tilde{\phi}_t^{k_t}(x_t) + (\tilde{H} \cdot \boldsymbol{x})_n \quad \textit{and}$$

$$\phi_t^k + \chi_t^k - \chi_{t+1} = \tilde{\phi}_t^k + \tilde{\chi}_t^k - \tilde{\chi}_{t+1} \quad \textit{for all } k \geq 1,\ t = 0, \dots, n,$$

as well as $\boldsymbol{\mu}(\boldsymbol{\phi}) = \boldsymbol{\mu}(\tilde{\boldsymbol{\phi}})$.

Proof. (i) If $\boldsymbol{x}$ is such that $(x_{t-1}, x_t) \notin I_k^t \times J_k^t$, then $\tilde{\phi}_t^{k_t}(x_t) = \phi_t^{k_t}(x_t)$ for $t = 0, \dots, n$ and $\tilde{H}(\boldsymbol{x}) = H(\boldsymbol{x})$. Otherwise,

$$\begin{aligned}
\tilde{\phi}_t^{k_t}(x_t) + \tilde{\phi}_{t-1}^{k_{t-1}}(x_{t-1}) + \tilde{H}_t(x_t - x_{t-1}) &= \phi_t^{k_t}(x_t) + \phi_{t-1}^{k_{t-1}}(x_{t-1}) \\
&\quad + H_t(x_t - x_{t-1}), \\
\tilde{\phi}_t^k + \tilde{\chi}_t^k - \tilde{\chi}_{t+1} &= \phi_t^k + \chi_t^k - \chi_{t+1}, \text{ and} \\
\tilde{\phi}_{t-1}^{k'} + \tilde{\chi}_{t-1}^{k'} - \tilde{\chi}_t &= \phi_{t-1}^{k'} + \chi_{t-1}^{k'} - \chi_t.
\end{aligned}$$

Along with the fact that $(\mu_t - \mu_{t-1})^k(\chi_t^k) = (\mu_t - \mu_{t-1})^k(\tilde{\chi}_t^k)$, these identities imply the assertions.

(ii) Similarly as in (i), the terms in question coincide by construction. □

Remark 1.4.14. The modification of Lemma 1.4.13 (i) can be applied simultaneously for infinitely many k's without difficulties. In this case we set

$$\tilde{\phi}_{t-1}^{k'}(x) := \phi_{t-1}^{k'}(x) + \sum_{k \geq 1} (c_1^k x - c_2^k)|_{I_k^t},$$

as well as $\tilde{\phi}_t^k(y) = \phi_t^k(y) - (c_1^k y - c_2^k)$ and $\tilde{\chi}_t^k(y) = \chi_t^k(y) + (c_1^k y - c_2^k)$ for the components $k \geq 1$ in step t. The pointwise equalities still hold as above and in particular, the moderated integral does not change.

Remark 1.4.15. Any $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)$ induces an element $(\boldsymbol{\phi}^g, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)$ with $\boldsymbol{\mu}(\boldsymbol{\phi}^g) = \boldsymbol{\mu}(\boldsymbol{\phi})$ by choosing some concave moderator $\boldsymbol{\chi}$ for $\boldsymbol{\phi}$ and setting

$$\phi_t^k := \phi_t|_{J_k^t}, \qquad \chi_t^k := \chi_t|_{J_k^t}.$$

We now show the analogue to Lemma 1.4.8 for the generalized dual space.

Lemma 1.4.16. *Let $\boldsymbol{\phi} \in L^{c,g}(\boldsymbol{\mu})$ and let $H = (H_1, \dots, H_n)$ be $\mathbb{F}$-predictable. If*

$$\sum_{t=0}^{n} \phi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t) + (H \cdot \boldsymbol{x})_n$$

is bounded from below on the effective domain $\mathcal{V}$ of $\mathcal{M}(\boldsymbol{\mu})$, then

$$\boldsymbol{\mu}(\phi) = P\left[\sum_{t=0}^{n} \phi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t) + (H \cdot \boldsymbol{x})_n\right], \quad P \in \mathcal{M}(\boldsymbol{\mu}).$$

Proof. Let $P \in \mathcal{M}(\boldsymbol{\mu})$, let $\boldsymbol{\chi}$ be a concave moderator for $\boldsymbol{\phi}$ such that $\chi_t^0 \equiv 0$ and assume that 0 is the lower bound. It is easy to see that $\sum_{t=0}^{n} \phi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t) + (H \cdot \boldsymbol{x})_n$ equals

$$\sum_{t=0}^{n}(\phi_t^{\boldsymbol{k}_t(\boldsymbol{x})} - \chi_{t+1} + \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})})(x_t) + \sum_{t=1}^{n}(\chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_{t-1}) - \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t)) + (H \cdot \boldsymbol{x})_n \geq 0.$$

By assumption $\sum_{t=0}^{n}(\phi_t^{\boldsymbol{k}_t(\boldsymbol{x})} - \chi_{t+1} + \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})})(x_t)$ is P-integrable. Therefore, the negative part of the remaining expression must also be P-integrable. Writing $P_t := P \circ (X_0, \dots, X_t)^{-1}$ and using that $(\chi_t^k)^+$ has linear growth, we see that for any disintegration $P = P_{n-1} \otimes \kappa_n$,

$$\begin{aligned}
&\int \left[\sum_{t=1}^{n}(\chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_{t-1}) - \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t)) + (H \cdot \boldsymbol{x})_n\right] \kappa_n(x_0, \dots, x_{n-1}, dx_n) \\
&= \sum_{t=1}^{n-1}(\chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_{t-1}) - \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t)) + (H \cdot \boldsymbol{x})_{n-1} \\
&\qquad + \int \left[(\chi_n^{\boldsymbol{k}_n(\boldsymbol{x})}(x_{n-1}) - \chi_n^{\boldsymbol{k}_n(\boldsymbol{x})}(x_n))\right] \kappa_n(x_0, \dots, x_{n-1}, dx_n).
\end{aligned}$$

Iteratively integrating with kernels such that $P_t = P_{t-1} \otimes \kappa_t$ and observing that we can apply Fubini's theorem to $\sum_{t=1}^{n}(\chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_{t-1}) - \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t)) + (H \cdot \boldsymbol{x})_n$ as its negative part is P-integrable, we obtain

$$P\left[\sum_{t=1}^{n}(\chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_{t-1}) - \chi_t^{\boldsymbol{k}_t(\boldsymbol{x})}(x_t)) + (H \cdot \boldsymbol{x})_n\right] = \sum_{t=1}^{n}\sum_{k \geq 1}(\mu_{t-1} - \mu_t)^k(\chi_t^k)$$

and the result follows. $\square$

Next, we establish that lifting from $\mathcal{D}_{\boldsymbol{\mu}}(f)$ to $\mathcal{D}^g_{\boldsymbol{\mu}}(f)$ does not change the range of dual costs.

Proposition 1.4.17. *Let $f : \mathbb{R}^{n+1} \to [0, \infty]$. We have*

$$\{\boldsymbol{\mu}(\boldsymbol{\phi}^g) : (\boldsymbol{\phi}^g, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)\} = \{\boldsymbol{\mu}(\boldsymbol{\phi}) : (\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)\}.$$

Proof. Remark 1.4.15 shows the inclusion "$\supseteq$." To show the reverse, we may apply Lemma 1.4.13 (i) together with Remark 1.4.14 to modify a given pair $(\boldsymbol{\phi}^g, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)$ such that $\phi^k_t(x) = 0$ for $x \in J^t_k \backslash I^t_k$, for all irreducible domains (I^t_k, J^t_k) of $\mathcal{M}(\mu_{t-1}, \mu_t)$ and $1 \leq t \leq n$. Here we have used that $x \in J^t_k \backslash I^t_k$ implies $\mu^t_k(\{x\}) > 0$, cf. Definition 1.2.2, and therefore $\phi^g \in L^{c,g}(\boldsymbol{\mu})$ implies $\phi^k_t(x) \in \mathbb{R}$; that is, such endpoints can indeed be shifted to 0 by adding affine functions to ϕ^k_t.

Let $\boldsymbol{\chi}^g$ be a concave moderator for $\boldsymbol{\phi}^g$. Using Lemma 1.4.3 (ii) and again Lemma 1.4.13 as above, we can modify χ^k_t to satisfy $\chi^k_t(x) = 0$ for $x \in J^t_k \backslash I^t_k$, for all irreducible domains (I^t_k, J^t_k) of $\mathcal{M}(\mu_{t-1}, \mu_t)$ and $1 \leq t \leq n$. Here, the finiteness of χ^k_t at the endpoints follows from Lemma 1.4.3 (i) and $(\mu_{t-1} - \mu_t)^k(\chi^k_t) < \infty$.

Still denoting the modified dual element by $(\boldsymbol{\phi}^g, H)$, we define $\boldsymbol{\phi} \in L^c(\boldsymbol{\mu})$ and a corresponding concave moderator $\boldsymbol{\chi}$ by

$$\phi_t(x) := \phi^k_t(x), \quad \chi_t(x) := \chi^k_t(x), \qquad \text{for } x \in J^t_k;$$

they are well-defined since ϕ^k_t and χ^k_t vanish at points that belong to more than one set J^t_k. We have $\boldsymbol{\mu}(\boldsymbol{\phi}) = \boldsymbol{\mu}(\boldsymbol{\phi}^g)$ by construction and the result follows. ☐

Definition 1.4.18. Let $1 \leq t \leq n$ and $x_t \in \mathbb{R}$. A sequence $\boldsymbol{x} = (x_0, \dots, x_t)$ is a *predecessor path* of x_t if there are indices $(k_0, \dots, k_t)$ such that $(x_{s-1}, x_s) \in (I^s_{k_s}, J^s_{k_s})$ for some component (irreducible or diagonal) of $\mathcal{M}(\mu_{s-1}, \mu_s)$, for all $1 \leq s \leq t$. We write $\Bbbk(\boldsymbol{x})$ for the (unique) associated sequence $(k_0, \dots, k_t)$ followed by the path $\boldsymbol{x}$ in the above sense, and $\Psi^k_t(x_t)$ for the set of all predecessor paths with $k_t = k$.

These notions will be useful in the next step towards the closedness result, which is to "regularize" the concave moderators. For concreteness in some of the expressions below, we convene that $\infty - \infty := \infty$.

Lemma 1.4.19. *Let $(\boldsymbol{\phi}, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(0)$. There is a concave moderator $\boldsymbol{\chi}$ of $\boldsymbol{\phi}$ such that*

$$\phi^k_t + \chi^k_t - \chi_{t+1} \geq 0 \quad \text{on } J^t_k \quad \text{for all} \quad t = 0, \dots, n, \quad k \geq 1, \text{ and} \tag{1.4.2}$$

$$\phi^0_t + \chi^0_t - \chi_{t+1} \geq 0 \quad \mu_t\text{-a.s. on } I^t_0 \quad \text{for all} \quad t = 1, \dots, n. \tag{1.4.3}$$

As a consequence,

$$\sum_{t=1}^{n} \sum_{k \geq 1} (\mu_{t-1} - \mu_t)^k (\chi^k_t) \leq \boldsymbol{\mu}(\boldsymbol{\phi}).$$

Proof. Fix $1 \leq t \leq n$ and let (I^t_k, J^t_k) be the domain of some component of $\mathcal{M}(\mu_{t-1}, \mu_t)$. We define $\boldsymbol{\chi} = (\chi^k_t)$ by $\chi^0_0 = 0$ and

$$\chi^k_t(x_t) = \inf_{\boldsymbol{x} \in \Psi^k_t(x_t)} \left\{ \sum_{s=0}^{t-1} \phi^{\mathbb{k}_s(\boldsymbol{x})}_s(x_s) + (H \cdot \boldsymbol{x})_t \right\};$$

then χ^k_t is concave on J^t_k for $k \geq 1$ as an infimum of affine functions.

We first show that

$$\{\chi^k_t = +\infty\} \subseteq \left\{\phi^{k'}_{t-1} = +\infty\right\} \cup \{\chi^{k'}_{t-1} = +\infty\}.$$

In particular, such points only exist after a chain of diagonal components from a point where $\phi^k_t(x_t) = \infty$. Suppose $\chi^k_t(x_t) = +\infty$ and $k \geq 1$, then the predecessor paths of x_t agree with the predecessor paths of all of J^k_t up to $t-1$, but $\{\sum_{s=0}^{t-1} \phi^{\mathbb{k}_s(\boldsymbol{x})}_s(x_s) < \infty\}$ must hold $\mathcal{M}(\boldsymbol{\mu})$-q.s. as $\boldsymbol{\phi} \in L^{c,g}(\boldsymbol{\mu})$. We must therefore have $x_t \in I^0_t$. Then, by definition, $\chi^0_t(x_t) = \chi^k_{t-1}(x_t) + \phi^k_{t-1}(x_t)$ and the claim follows.

Next, we verify that $\boldsymbol{\chi}$ satisfies (1.4.2) and (1.4.3). For notational convenience we for now set $\chi_{n+1} \equiv \inf_{\boldsymbol{x} \in \mathcal{V}} \left\{ \sum_{s=0}^{n} \phi^{\mathbb{k}_s(\boldsymbol{x})}_s(x_s) + (H \cdot \boldsymbol{x})_n \right\} \geq 0$. Restricting the infimum in the definition of $\boldsymbol{\chi}$ to the set of paths $\boldsymbol{x}$ with $x_{t+1} = x_t \in I_{k'}^{t+1} \cap J_k^t$ yields

$$
\begin{aligned}
\chi_{t+1}(x_t) = \chi_{t+1}^{k'}(x_t) &= \inf_{\boldsymbol{x} \in \Psi_{t+1}^{k'}(x_t)} \left\{ \sum_{s=0}^{t} \phi_s^{\Bbbk_s(\boldsymbol{x})}(x_s) + (H \cdot \boldsymbol{x})_{t+1} \right\} \\
&\leq \inf_{\boldsymbol{x} \in \Psi_t^k(x_t)} \left\{ \sum_{s=0}^{t-1} \phi_s^{\Bbbk_s(\boldsymbol{x})}(x_s) + (H \cdot \boldsymbol{x})_t \right\} + \phi_t^k(x_t) \\
&= \chi_t^k(x_t) + \phi_t^k(x_t).
\end{aligned}
$$

Since $\cup_{k' \geq 0} I_{k'}^{t+1} = \mathbb{R}$, this will imply (1.4.2) after we check that $\chi_t^k > -\infty$ for $k \geq 1$ and $\chi_t^0 > -\infty$ holds μ_t^0-a.s., which also implies that $\chi_t > -\infty$ holds μ_{t-1}-almost surely. We show this inductively for $t \geq 1$.

Clearly $\chi_{n+1} \geq 0 > -\infty$. Now, for $t \leq n$ the induction hypothesis is that $\chi_{t+1} > -\infty$ holds almost surely μ_t.

From $\boldsymbol{\phi} \in L^{c,g}$ and $\chi_{t+1} > -\infty$ μ_t-a.s. we have that

$$
\phi_t^k < \infty, \quad \chi_{t+1} > -\infty \quad \text{hold } \mu_t^k\text{-a.s.}
$$

As χ_t^k is concave and J_t^k is the convex hull of the topological support of μ_t^k we then get $\chi_t^k > -\infty$ on all of J_t^k from the previous inequality.

For $k = 0$, the inequality yields $\{\chi_t^0 = -\infty\} \subseteq \{\chi_{t+1} = -\infty\} \cup \{\phi_t^0(x_t) = \infty\}$ and both of these sets are μ_t nullsets. Finally $\mu_{t-1}(\{\chi_t = -\infty\}) = 0$ as this is a subset of the diagonal component where μ_{t-1} is dominated by μ_t.

Set $\bar{\phi}_t^k := \phi_t^k + \chi_t^k - \chi_{t+1}|_{J_k^t}$ for $0 \leq t \leq n$; then $\bar{\phi}_t^k \geq 0$. Moreover, choose an arbitrary $P \in \mathcal{M}(\boldsymbol{\mu})$ with disintegration $P = \mu_0 \otimes \kappa_1 \otimes \cdots \otimes \kappa_n$ for some stochastic kernels $\kappa_t(x_0, \ldots, x_{t-1}, dx_t)$. From Lemma 1.4.16 we know that

$$
\boldsymbol{\mu}(\boldsymbol{\phi}) = P\left[\sum_{t=0}^{n} \phi_t^{\Bbbk_t(X)}(X_t) + (H \cdot X)_n \right] < \infty.
$$

We can therefore apply Fubini's theorem for kernels as in the proof of Lemma 1.4.16 to the expression

$$
\begin{aligned}
0 &\leq \sum_{t=0}^{n} \phi_t^{\Bbbk_t(\boldsymbol{x})}(x_t) + (H \cdot \boldsymbol{x})_n \\
&= \sum_{t=0}^{n} \bar{\phi}_t^{\Bbbk_t(\boldsymbol{x})}(x_t) + \sum_{t=1}^{n} \Big(\chi_t(x_{t-1}) - \chi_t^{\Bbbk_t(\boldsymbol{x})}(x_t) \Big) + (H \cdot \boldsymbol{x})_n
\end{aligned}
$$

and obtain

$$
P\left[\sum_{t=0}^{n} \phi_t^{\Bbbk_t(X)}(X_t) + (H \cdot X)_n \right] = \sum_{t=0}^{n} \sum_{k \geq 0} \mu_t^k(\bar{\phi}_t^k) + \sum_{t=1}^{n} \sum_{k \geq 1} (\mu_{t-1} - \mu_t)^k (\chi_t^k)
$$

which shows that the right hand side is finite, and therefore $\boldsymbol{\chi}$ is a concave moderator for $\boldsymbol{\phi}$. Finally, the second claim follows from $\mu_t^k(\bar{\phi}_t^k) \geq 0$. $\square$

The last tool for our closedness result is a compactness property for concave functions in the one-step case; cf. [18, Proposition 5.5].

Proposition 1.4.20. *Let $\mu \leq_c \nu$ be irreducible with domain (I, J) and let $a \in I$ be the common barycenter of μ and ν. Let $\chi_m : J \to \mathbb{R}$ be concave functions such that*[8]

$$
\chi_m(a) = \chi_m'(a) = 0 \qquad \textit{and} \qquad \sup_{m \geq 1} (\mu - \nu)(\chi_m) < \infty.
$$

There exists a subsequence χ_{m_k} which converges pointwise on J to a concave function $\chi : J \to \mathbb{R}$, and $(\mu - \nu)(\chi) \leq \liminf_k (\mu - \nu)(\chi_{m_k})$.

We are now ready to state and prove the analogue of Proposition 1.4.10 in the generalized dual.

Proposition 1.4.21. *Let $f^m : \mathbb{R}^{n+1} \to [0, \infty]$, $m \geq 1$ be a sequence of functions such that*

$$
f^m \to f \qquad \textit{pointwise}
$$

and let $(\boldsymbol{\phi}^m, H^m) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f^m)$ be such that $\sup_m \boldsymbol{\mu}(\boldsymbol{\phi}^m) < \infty$. Then there exist $(\boldsymbol{\phi}, H) \in$

[8] To be specific, let us convene that χ_m' is the left derivative—this is not important here.

$\mathcal{D}^g_{\boldsymbol{\mu}}(f)$ *with*

$$\boldsymbol{\mu}(\boldsymbol{\phi}) \leq \liminf_{m\to\infty} \boldsymbol{\mu}(\boldsymbol{\phi}^m).$$

Proof. Since $(\boldsymbol{\phi}^m, H^m) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f^m)$ and $f^m \geq 0$, we can introduce a sequence of concave moderators $\boldsymbol{\chi}_m$ as in Lemma 1.4.19. A normalization of $(\boldsymbol{\phi}^m, H^m)$ as in Lemma 1.4.13 (i) and (ii), in the general form of Remark 1.4.14, allows us to assume without loss of generality that $\chi^0_{t,m} \equiv 0$ and $\chi^k_{t,m}(a^k_t) = (\chi^k_{t,m})'(a^k_t) = 0$, where a^k_t is the barycenter of μ^k_t—this modification is the main merit of lifting to the generalized dual space. While the generalized dual gives enough degrees of freedom to choose this normalization, the dual without the generalization does not. This is related to the possible overlap of the intervals I, J at the different times t; see also Figure 1.2 and the paragraph preceding Example 1.3.2.

By passing to a subsequence as in Proposition 1.4.20 for each component and using a diagonal argument, we obtain pointwise limits $\chi^k_t : J^t_k \to \mathbb{R}$ for $\chi^k_{t,m}$ after passing to another subsequence.

Since $\phi^k_{t,m} + \chi^k_{t,m} - \chi_{t+1,m} \geq 0$ on J^t_k[9] and $\chi^k_{t,m} \to \chi^k_t$ as well as $\chi_{t+1,m} \to \chi_{t+1}$, we can apply Komlos' lemma (in the form of [44, Lemma A1.1] and its remark) to find convex combinations $\tilde{\phi}^k_{t,m} \in \operatorname{conv}\{\phi^k_{t,m}, \phi^k_{t,m+1}, \dots\}$ which converge μ^k_t-a.s. for $0 \leq t \leq n$. We may assume without loss of generality that $\tilde{\phi}^k_{t,m} = \phi^k_{t,m}$. Thus, we can set

$$\begin{aligned} \phi^k_t &:= \limsup \phi^k_{t,m} \quad \text{ on } J^t_k \quad \text{for } t = 1, \dots, n, \\ \phi_0 &:= \liminf \phi_{0,m} \end{aligned}$$

to obtain

$$\phi^k_{t,m} \to \phi^k_t \quad \mu^k_t\text{-a.s.} \quad \text{ and } \quad \phi^k_t + \chi^k_t - \chi_{t+1} \geq 0 \text{ on } J^t_k.$$

[9] Observe that this inequality will still hold after modifying $\boldsymbol{\phi}$ and $\boldsymbol{\chi}$ as in Lemma 1.4.13.

We can now apply Fatou's lemma and Proposition 1.4.20 to deduce that

$$\begin{aligned}
\boldsymbol{\mu}(\boldsymbol{\phi}) &= \sum_{t=0}^{n}\sum_{k\geq 0}\mu_t^k(\phi_t^k+\chi_t^k-\chi_{t+1}) + \sum_{t=1}^{n}\sum_{k\geq 1}(\mu_{t-1}-\mu_t)^k(\chi_t^k) \\
&\leq \sum_{t=0}^{n}\sum_{k\geq 0}\liminf \mu_t^k(\phi_{t,m}^k+\chi_{t,m}^k-\chi_{t+1,m}) \\
&\quad + \sum_{t=1}^{n}\sum_{k\geq 1}\liminf(\mu_{t-1}-\mu_t)^k(\chi_{t,m}^k) \\
&\leq \liminf\left[\sum_{t=0}^{n}\sum_{k\geq 0}\mu_t^k(\phi_{t,m}^k+\chi_{t,m}^k-\chi_{t+1,m}) + \sum_{t=1}^{n}\sum_{k\geq 1}(\mu_{t-1}-\mu_t)^k(\chi_{t,m}^k)\right] \\
&= \liminf \boldsymbol{\mu}(\boldsymbol{\phi}^m) < \infty.
\end{aligned}$$

In particular, we see that $\boldsymbol{\phi} \in L^{c,g}(\boldsymbol{\mu})$ with concave moderator $\boldsymbol{\chi}$.

It remains to construct the predictable process $H = (H_1, \dots, H_n)$. With a mild abuse of notation, we shall identify $H_t(x_0, \dots, x_n)$ with the corresponding function of $(x_0, \dots, x_{t-1})$ in this proof.

We first define for each $\boldsymbol{k} = (k_0, \dots, k_t)$ and $\boldsymbol{x} = (x_0, \dots, x_t)$ such that $\boldsymbol{k} = \Bbbk(\boldsymbol{x})$, the functions $G_{t,m}^{\boldsymbol{k}}$ and $G_t^{\boldsymbol{k}}$ by

$$\begin{aligned}
G_{t,m}^{\boldsymbol{k}}(\boldsymbol{x}) &:= \sum_{s=0}^{t}\phi_{s,m}^{k_s}(x_s) + \sum_{s=1}^{t}H_{s,m}(x_0, \dots, x_{s-1})\cdot(x_s - x_{s-1}), \\
G_t^{\boldsymbol{k}}(\boldsymbol{x}) &:= \liminf G_{t,m}^{\boldsymbol{k}}(\boldsymbol{x}).
\end{aligned}$$

Given $\boldsymbol{k} = (k_0, \dots, k_t)$, we write $\boldsymbol{k}' = (k_0, \dots, k_{t-1})$. We claim that there exists an $\mathbb{F}$-predictable process H such that for all $1 \leq t \leq n$,

$$G_{t-1}^{\boldsymbol{k}'}(x_0, \dots, x_{t-1}) + \phi_t^{k_t}(x_t) + H_t(x_0, \dots, x_{t-1})\cdot(x_t - x_{t-1}) \geq G_t^{\boldsymbol{k}}(x_0, \dots, x_t). \quad (1.4.4)$$

Once this is established, the proposition follows by induction since $G_0^{(0)}(x_0) = \phi_0(x_0)$ and $G_n^{\boldsymbol{k}}(x_0, \dots, x_n) \geq f(x_0, \dots, x_n)$.

To prove the claim, write g^{conc} for the concave hull of a function g and observe that

$$\begin{aligned}
&\liminf[G^{\boldsymbol{k}'}_{t-1,m}(x_0,\dots,x_{t-1}) + H_{t,m}(x_0,\dots,x_{t-1})\cdot(x_t - x_{t-1})]\\
&\qquad\qquad \geq \liminf[(G^{\boldsymbol{k}}_{t,m}(x_0,\dots,x_{t-1},\cdot) - \phi^{k_t}_{t,m}(\cdot))^{\mathrm{conc}}(x_t)]\\
&\qquad\qquad \geq [\liminf(G^{\boldsymbol{k}}_{t,m}(x_0,\dots,x_{t-1},\cdot) - \phi^{k_t}_{t,m}(\cdot)]^{\mathrm{conc}}(x_t)\\
&\qquad\qquad \geq [G^{\boldsymbol{k}}_{t}(x_0,\dots,x_{t-1},\cdot) - \phi^{k_t}_{t}(\cdot)]^{\mathrm{conc}}(x_t)\\
&\qquad\qquad =: \hat{\phi}^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1},x_t).
\end{aligned}$$

By construction, $\hat{\phi}^{\boldsymbol{k}}_t$ is concave in the last variable and satisfies

$$G^{\boldsymbol{k}'}_{t-1}(x_0,\dots,x_{t-1}) \geq \hat{\phi}^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1},x_{t-1}).$$

Let $\partial_t\hat{\phi}^{\boldsymbol{k}}_t$ denote the left partial derivative in the last variable and set

$$H^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1}) := \partial_t\hat{\phi}^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1},x_{t-1})$$

for $\boldsymbol{k}_t \geq 1$ and $H^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1}) = 0$ for $\boldsymbol{k}_t = 0$; then we have

$$\begin{aligned}
&G^{\boldsymbol{k}'}_{t-1}(x_0,\dots,x_{t-1}) + H^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1})\cdot(x_t - x_{t-1})\\
&\qquad \geq \hat{\phi}^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1},x_{t-1}) + H^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1})\cdot(x_t - x_{t-1})\\
&\qquad \geq \hat{\phi}^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1},x_t)\\
&\qquad \geq G^{\boldsymbol{k}}_t(x_0,\dots,x_t) - \phi^{k_t}_t(x_t).
\end{aligned}$$

Finally, for any $(x_0,\dots,x_{t-1}) \in \mathbb{R}^t$, we define $H_t(x_0,\dots,x_{t-1})$ as

$$\begin{cases} H^{\boldsymbol{k}}_t(x_0,\dots,x_{t-1}), & \text{if } \boldsymbol{k} = \Bbbk(x_0,\dots,x_{t-1},x_t) \text{ for some } x_t \in \mathbb{R}\\ 0, & \text{otherwise;}\end{cases}$$

this is well-defined since $\Bbbk(x_0,\dots,x_t)$ depends only on $(x_0,\dots,x_{t-1})$. The predictable process H satisfies (1.4.4) and thus the proof is complete. □

Proof of Proposition 1.4.10. In view of Remark 1.4.15 and Proposition 1.4.17, the result follows from Proposition 1.4.21. □

1.5 Duality Theorem and Monotonicity Principle

The first goal of this section is a duality result for the multistep martingale transport problem; it establishes the absence of a duality gap and the existence of optimizers in the dual problem. (As is well known, an optimizer for the primal problem only exists under additional conditions, such as continuity of f.) The second goal is a monotonicity principle describing the geometry of optimal transports; it will be a consequence of the duality result.

As above, we consider a fixed vector $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ of marginals in convex order. The primal and dual problems as defined follows.

Definition 1.5.1. Let $f : \mathbb{R}^{n+1} \to [0, \infty]$. The *primal problem* is

$$\mathbf{S}_{\boldsymbol{\mu}}(f) := \sup_{P \in \mathcal{M}(\boldsymbol{\mu})} P(f) \in [0, \infty],$$

where $P(f)$ refers to the outer integral if f is not measurable. The *dual problem* is

$$\mathbf{I}_{\boldsymbol{\mu}}(f) := \inf_{(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)} \boldsymbol{\mu}(\boldsymbol{\phi}) \in [0, \infty].$$

We recall that a function $f : \mathbb{R}^{n+1} \to [0, \infty]$ is called *upper semianalytic* if the sets $\{f \geq c\}$ are analytic for all $c \in \mathbb{R}$, where a subset of $\mathbb{R}^{n+1}$ is called analytic if it is the image of a Borel subset of a Polish space under a Borel mapping. Any Borel function is upper semianalytic and any upper semianalytic function is universally measurable; we refer to [20, Chapter 7] for background. The following is the announced duality result.

Theorem 1.5.2 (Duality)**.** *Let* $f : \mathbb{R}^{n+1} \to [0, \infty]$.

(i) If f is upper semianalytic, then $\mathbf{S}_{\boldsymbol{\mu}}(f) = \mathbf{I}_{\boldsymbol{\mu}}(f) \in [0, \infty]$.

(ii) If $\mathbf{I}_{\boldsymbol{\mu}}(f) < \infty$*, there exists a dual optimizer* $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)$.

Proof. Given our preceding results, much of the proof follows the lines of the corresponding result for the one-step case in [18, Theorem 6.2]; therefore, we shall be brief. We mention that the present theorem is slightly more general than the cited one in terms of the measurability condition (f is upper semianalytic instead of Borel); this is due to the global proof given here.

Step 1. Using Lemma 1.4.8 we see that $\mathbf{S}_{\boldsymbol{\mu}}(f) \leq \mathbf{I}_{\boldsymbol{\mu}}(f)$ holds for all upper semicontinuous $f : \mathbb{R}^{n+1} \to [0, \infty]$.

Step 2. Using the de la Vallée–Poussin theorem and our assumption that the marginals have a finite first moment, there exist increasing, superlinearly growing functions $\zeta_{\mu_t} : \mathbb{R}^+ \to \mathbb{R}^+$ such that $x \mapsto \zeta_{\mu_t}(|x|)$ is μ_t-integrable for all $0 \leq t \leq n$. Define

$$\zeta(x_0, \dots, x_n) := 1 + \sum_{t=0}^{n} \zeta_{\mu_t}(|x_t|)$$

and let C_ζ be the vector space of all continuous functions f such that f/ζ vanishes at infinity. Then, a Hahn–Banach separation argument can be used to show that $\mathbf{S}_{\boldsymbol{\mu}}(f) \geq \mathbf{I}_{\boldsymbol{\mu}}(f)$ holds for all $f \in C_\zeta$; the details of the argument are the same as in the proof of [18, Lemma 6.4].

Step 3. Let f be bounded and upper semicontinuous; then there exists a sequence of bounded continuous functions $f^m \in C_b(\mathbb{R}^{n+1})$ which decrease to f pointwise. As $C_b(\mathbb{R}^{n+1}) \subseteq C_\zeta$, we have $\mathbf{S}_{\boldsymbol{\mu}}(f^m) = \mathbf{I}_{\boldsymbol{\mu}}(f^m)$ for all m by the first two steps.

Let $\mathcal{U}$ be the set of all bounded, nonnegative, upper semicontinuous functions on $\mathbb{R}^{n+1}$. We recall that a map $\mathbf{C} : [0, \infty]^{\mathbb{R}^{n+1}} \to [0, \infty]$ is called a $\mathcal{U}$-capacity if it is monotone, sequentially continuous upwards on $[0, \infty]^{\mathbb{R}^{n+1}}$ and sequentially continuous downwards on $\mathcal{U}$. The functional $f \mapsto \mathbf{S}_{\boldsymbol{\mu}}(f)$ is a $\mathcal{U}$-capacity; this follows from the weak compactness of $\mathcal{M}(\boldsymbol{\mu})$ and the arguments in [78, Propositions 1.21, 1.26].

It follows that $\mathbf{S}_{\boldsymbol{\mu}}(f^m) \to \mathbf{S}_{\boldsymbol{\mu}}(f)$. By the monotonicity of $f \mapsto \mathbf{I}_{\boldsymbol{\mu}}(f)$ and Step 1 we obtain

$$\mathbf{I}_{\boldsymbol{\mu}}(f) \leq \lim \mathbf{I}_{\boldsymbol{\mu}}(f^m) = \lim \mathbf{S}_{\boldsymbol{\mu}}(f^m) = \mathbf{S}_{\boldsymbol{\mu}}(f) \leq \mathbf{I}_{\boldsymbol{\mu}}(f).$$

Step 4. Since $\mathbf{S}_{\boldsymbol{\mu}} = \mathbf{I}_{\boldsymbol{\mu}}$ on $\mathcal{U}$ by Step 3, $\mathbf{I}_{\boldsymbol{\mu}}$ is sequentially downward continuous on $\mathcal{U}$ like $\mathbf{S}_{\boldsymbol{\mu}}$. On the other hand, Proposition 1.4.10 implies that it is sequentially upwards continuous on $[0, \infty]^{\mathbb{R}^{n+1}}$. As a result, $\mathbf{I}_{\boldsymbol{\mu}}$ is a $\mathcal{U}$-capacity.

Step 5. Let $f : \mathbb{R}^{n+1} \to [0, \infty]$ be upper semianalytic. For any $\mathcal{U}$-capacity $\mathbf{C}$, Choquet's capacitability theorem shows that

$$\mathbf{C}(f) = \sup\{\mathbf{C}(g) : g \in \mathcal{U}, g \leq f\}.$$

As $\mathbf{S}_{\boldsymbol{\mu}}$ and $\mathbf{I}_{\boldsymbol{\mu}}$ are $\mathcal{U}$-capacities that coincide on $\mathcal{U}$, it follows that $\mathbf{S}_{\boldsymbol{\mu}}(f) = \mathbf{I}_{\boldsymbol{\mu}}(f)$. This completes the proof of (i).

Step 6. To see that the infimum $\mathbf{I}_{\boldsymbol{\mu}}(f)$ is attained if it is finite, we merely need to apply Proposition 1.4.10 with the constant sequence $f^m = f$. □

We can easily relax the lower bound on f.

Remark 1.5.3. Let $f : \mathbb{R}^{n+1} \to (-\infty, \infty]$ and suppose there exist $\boldsymbol{\phi} \in \prod_{t=0}^{n} L^1(\mu_t)$ and a predictable process H such that

$$f \geq \sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n \quad \text{on} \quad \mathcal{V}.$$

Then we can apply Theorem 1.5.2 to $[f - \sum_{t=0}^{n} \phi_t(X_t) - (H \cdot X)_n]^+$ and obtain the analogue of its assertion for f.

The duality result gives rise to a monotonicity principle describing the support of optimal martingale transports, in the spirit of the cyclical monotonicity condition from classical transport theory. The following generalizes the results of [16, Lemma 1.11] and [18, Corollary 7.8] for the one-step martingale transport problem.

Theorem 1.5.4 (Monotonicity Principle)**.** *Let $f : \mathbb{R}^{n+1} \to [0, \infty]$ be Borel and suppose that $\mathbf{S}_{\boldsymbol{\mu}}(f) < \infty$. There exists a Borel set $\Gamma \subseteq \mathbb{R}^{n+1}$ with the following properties.*

(i) A measure $P \in \mathcal{M}(\boldsymbol{\mu})$ is concentrated on Γ if and only if it is optimal for $\mathbf{S}_{\boldsymbol{\mu}}(f)$.

(ii) Let $\bar{\boldsymbol{\mu}} = (\bar{\mu}_0, \ldots, \bar{\mu}_n)$ be another vector of marginals in convex order. If $\bar{P} \in \mathcal{M}(\bar{\boldsymbol{\mu}})$ is concentrated on Γ, then $\bar{P}$ is optimal for $\mathbf{S}_{\bar{\boldsymbol{\mu}}}(f)$.

Indeed, if $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)$ is an optimizer for $\mathbf{I}_{\boldsymbol{\mu}}(f)$, then we can take

$$\Gamma := \left\{ \sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n = f \right\} \cap \mathcal{V}.$$

Proof. As $\mathbf{S}_{\boldsymbol{\mu}}(f) < \infty$, Theorem 1.5.2 shows that $\mathbf{I}_{\boldsymbol{\mu}}(f) = \mathbf{S}_{\boldsymbol{\mu}}(f) < \infty$ and that there exists a dual optimizer $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)$. In particular, we can define Γ as above.

(i) As $0 \leq f$ and $P(f) \leq \mathbf{S}_{\boldsymbol{\mu}}(f) < \infty$ for all $P \in \mathcal{M}(\boldsymbol{\mu})$, we see that f is P-integrable for all $P \in \mathcal{M}(\boldsymbol{\mu})$. Since $\sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n \geq 0$ on the effective domain $\mathcal{V}$, and $P\left[\sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n\right] = \boldsymbol{\mu}(\boldsymbol{\phi}) = \mathbf{I}_{\boldsymbol{\mu}}(f) < \infty$ by Lemma 1.4.8, we also obtain the P-integrability of $\sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n$. In particular,

$$0 \leq P\left[\sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n - f\right] = \boldsymbol{\mu}(\boldsymbol{\phi}) - P(f) = \mathbf{S}_{\boldsymbol{\mu}}(f) - P(f)$$

and equality holds if and only if P is concentrated on Γ.

(ii) We may assume that $\bar{P}$ is a probability measure with $\bar{P}(f) < \infty$. As a first step, we show that the effective domain $\bar{\mathcal{V}}$ of $\mathcal{M}(\bar{\boldsymbol{\mu}})$ is a subset of the effective domain $\mathcal{V}$ of $\mathcal{M}(\boldsymbol{\mu})$. To that end, it is sufficient to show that if $1 \leq t \leq n$ and $x \in \mathbb{R}$ are such that $u_{\mu_{t-1}}(x) = u_{\mu_t}(x)$, then $u_{\bar{\mu}_{t-1}}(x) = u_{\bar{\mu}_t}(x)$, and if moreover $\partial^+ u_{\mu_{t-1}}(x) = \partial^+ u_{\mu_t}(x)$, then $\partial^+ u_{\bar{\mu}_{t-1}}(x) = \partial^+ u_{\bar{\mu}_t}(x)$, and similarly for the left derivative ∂^- (cf. Proposition 1.2.3). Indeed, for t and x such that $u_{\mu_{t-1}}(x) = u_{\mu_t}(x)$, our assumption that $\Gamma \subseteq \mathcal{V}$ implies

$$\Gamma \subseteq (X_{t-1}, X_t)^{-1}\big((-\infty, x]^2 \cup [x, \infty)^2\big).$$

Using also that $\mathbb{E}^{\bar{P}}[X_t|\mathfrak{F}_{t-1}] = X_{t-1}$ and that $\bar{P}$ is concentrated on Γ,

$$\begin{aligned} u_{\bar{\mu}_{t-1}}(x) &= \mathbb{E}^{\bar{P}}[|X_{t-1} - x|] \\ &= \mathbb{E}^{\bar{P}}[(X_{t-1} - x)\mathbf{1}_{X_{t-1} \geq x}] + \mathbb{E}^{\bar{P}}[(x - X_{t-1})\mathbf{1}_{X_{t-1} \leq x}] \\ &= \mathbb{E}^{\bar{P}}[(X_t - x)\mathbf{1}_{X_{t-1} \geq x}] + \mathbb{E}^{\bar{P}}[(x - X_t)\mathbf{1}_{X_{t-1} \leq x}] \\ &= \mathbb{E}^{\bar{P}}[|X_t - x|] = u_{\bar{\mu}_t}(x) \end{aligned}$$

as desired. If in addition $\partial^+ u_{\mu_{t-1}}(x) = \partial^+ u_{\mu_t}(x)$, then $\Gamma \subseteq \mathcal{V}$ implies

$$\Gamma \subseteq (X_{t-1}, X_t)^{-1}\big((-\infty, x]^2 \cup (x, \infty)^2\big).$$

As $\bar{P}$ is concentrated on Γ, it follows that

$$\begin{aligned}\partial^+ u_{\bar{\mu}_{t-1}}(x) &= \bar{P}[X_{t-1} \leq x] - \bar{P}[X_{t-1} > x] \\ &= \bar{P}[X_t \leq x] - \bar{P}[X_t > x] = \partial^+ u_{\bar{\mu}_t}(x)\end{aligned}$$

as desired. The same argument can be used for the left derivative and we have shown that $\bar{\mathcal{V}} \subseteq \mathcal{V}$.

In view of that inclusion, the inequality $\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n \geq f$ holds on $\bar{\mathcal{V}}$. Since $\bar{P}$ is concentrated on Γ,

$$\bar{P}\left[\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n\right] = \bar{P}(f) < \infty.$$

We may follow the arguments in the proof of Lemma 1.4.19 to construct a moderator $\boldsymbol{\chi}$ and establish that $(\boldsymbol{\phi}, H) \in \mathcal{D}^g_{\bar{\boldsymbol{\mu}}}(f)$, where we are implicitly using the embedding detailed in Remark 1.4.15. (Note that the proof of Lemma 1.4.19 uses the condition $(\boldsymbol{\phi}, H) \in \mathcal{D}^g_{\bar{\boldsymbol{\mu}}}(0)$ only to establish $\bar{P}\left[\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n\right] < \infty$. In the present situation the latter is known a priori and the condition is not needed.) Then, we can modify $\boldsymbol{\chi}$ as in the proof of Proposition 1.4.17 to see that $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\bar{\boldsymbol{\mu}}}(f)$. As a result, we may apply Lemma 1.4.8 to obtain that

$$\bar{P}(f) = \bar{P}\left[\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n\right] = \bar{\boldsymbol{\mu}}(\boldsymbol{\phi}),$$

whereas for any other $P' \in \mathcal{M}(\bar{\boldsymbol{\mu}})$ we have

$$P'(f) \leq P'\left[\sum_{t=0}^n \phi_t(X_t) + (H \cdot X)_n\right] = \bar{\boldsymbol{\mu}}(\boldsymbol{\phi}) = \bar{P}(f).$$

This shows that $\bar{P} \in \mathcal{M}(\bar{\boldsymbol{\mu}})$ is optimal. □

1.6 Left-Monotone Transports

In this section we define left-monotone transports through a shadow property and prove their existence.

1.6.1 Preliminaries

Before moving on to the n-step case, we recall the essential definitions and results regarding the one-step version of the left-monotone transport (also called the Left-Curtain coupling). The first notion is the so-called shadow, and it will be useful to define it for measures $\mu \leq_{pc} \nu$ in *positive convex order*, meaning that $\mu(\phi) \leq \nu(\phi)$ for any nonnegative convex function ϕ. Clearly, this order is weaker than the convex order $\mu \leq_c \nu$, and it is worth noting that μ may have a smaller mass than ν. The following is the result of [16, Lemma 4.6].

Lemma 1.6.1. *Let $\mu \leq_{pc} \nu$. Then the set*

$$[\![\mu;\nu]\!] := \{\theta : \mu \leq_c \theta \leq \nu\}$$

is non-empty and contains a unique least element $\mathcal{S}^\nu(\mu)$ for the convex order:

$$\mathcal{S}^\nu(\mu) \leq_c \theta \text{ for all } \theta \in [\![\mu;\nu]\!] .$$

The measure $\mathcal{S}^\nu(\mu)$ is called the shadow *of μ in ν.*

It will be useful to have the following picture in mind: if μ is a Dirac measure, its shadow in ν is a measure θ of equal mass and barycenter, chosen such as to have minimal variance subject to the constraint $\theta \leq \nu$.

The second notion is a class of reward functions.

Definition 1.6.2. A Borel function $f : \mathbb{R}^2 \to \mathbb{R}$ is called *second-order Spence–Mirrlees* if $y \mapsto f(x', y) - f(x, y)$ is strictly convex for any $x < x'$.

We note that if f is sufficiently differentiable, this can be expressed as the cross-derivative condition $f_{xyy} > 0$ which has also been called the martingale Spence–Mirrlees condition, in analogy to the classical Spence–Mirrlees condition $f_{xy} > 0$.

In the one-step case, the left-monotone transport is unique and can be characterized as follows; cf. [16, Theorems 4.18, 4.21, 6.1] where this transport is called the Left-Curtain coupling, as well as [95, Theorem 1.2] for the third equivalence in the stated generality.

Proposition 1.6.3. *Let $\mu \leq_c \nu$ and $P \in \mathcal{M}(\mu, \nu)$. The following are equivalent:*

(i) For all $x \in \mathbb{R}$ and $A \in \mathfrak{B}(\mathbb{R})$,

$$P[(-\infty, x] \times A] = \mathcal{S}^{\nu}(\mu|_{(-\infty,x]})(A).$$

(ii) P is concentrated on a Borel set $\Gamma \subseteq \mathbb{R}^2$ satisfying

$$(x, y^-), (x, y^+), (x', y') \in \Gamma, \ \ x < x' \quad \Rightarrow \quad y' \notin (y^-, y^+).$$

(iii) P is an optimizer of $\mathbf{S}_{\mu,\nu}(f)$ for some (and then all) $f : \mathbb{R}^2 \to \mathbb{R}$ second-order Spence–Mirrlees such that there exist functions $a \in L^1(\mu)$, $b \in L^1(\nu)$ with $|f(x, y)| \le a(x) + b(y)$.

There exists a unique measure $\bar{P} \in \mathcal{M}(\mu, \nu)$ satisfying (i)–(iii), and $\bar{P}$ is called the (one-step) left-monotone transport.

If μ is a discrete measure, the characterization in (i) can be understood as follows: the left-monotone transport $\bar{P}$ processes the atoms of μ from left to right, mapping each one of them to its shadow in the remaining target measure.

Next, we record two more results about shadows that will be used below. The first one, cited from [17, Theorem 3.1], generalizes the above idea in the sense that the atoms are still mapped to their shadows but can be processed in any given order; in the general (non-discrete) case, such an order is defined by a coupling π from the uniform measure to μ.

Proposition 1.6.4. *Let $\mu \le_c \nu$ and $\pi \in \Pi(\lambda, \mu)$ where λ denotes the Lebesgue measure on $[0, 1]$. Then there exists a unique measure $Q \in \Pi(\lambda, \mu, \nu)$ on $\mathbb{R}^3$ such that $Q \circ (X_0, X_1)^{-1} = \pi$ and*

$$Q|_{[0,s]\times\mathbb{R}\times\mathbb{R}} \circ (X_1, X_2)^{-1} \in \mathcal{M}(\pi_s, \mathcal{S}^{\nu}(\pi_s)), \quad s \in \mathbb{R},$$

where $\pi_s := \pi|_{[0,s]\times\mathbb{R}} \circ (X_1)^{-1}$.

We shall also need the following facts about shadows.

Lemma 1.6.5. *(i) Let μ_1, μ_2, ν be finite measures satisfying $\mu_1 + \mu_2 \le_{pc} \nu$. Then $\mu_2 \le_{pc} \nu - \mathcal{S}^{\nu}(\mu_1)$ and $\mathcal{S}^{\nu}(\mu_1 + \mu_2) = \mathcal{S}^{\nu}(\mu_1) + \mathcal{S}^{\nu - \mathcal{S}^{\nu}(\mu_1)}(\mu_2)$.*

(ii) Let μ, ν_1, ν_2 be finite measures such that $\mu \leq_{pc} \nu_1 \leq_c \nu_2$. Then, it follows that $\mathcal{S}^{\nu_1}(\mu) \leq_{pc} \nu_2$. Moreover, $\mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_1}(\mu)) = \mathcal{S}^{\nu_2}(\mu)$ if and only if $\mathcal{S}^{\nu_1}(\mu) \leq_c \mathcal{S}^{\nu_2}(\mu)$.

Proof. Part (i) is [16, Theorem 4.8]. To obtain the first statement in (ii), we observe that $\mathcal{S}^{\nu_1}(\mu) \leq \nu_1 \leq_c \nu_2$ and hence

$$\mathcal{S}^{\nu_1}(\mu)(\phi) \leq \nu_1(\phi) \leq \nu_2(\phi)$$

for any nonnegative convex function ϕ. Turning to the second statement, the "only if" implication follows directly from the definition of the shadow in Lemma 1.6.1. To show the reverse implication, suppose that $\mathcal{S}^{\nu_1}(\mu) \leq_c \mathcal{S}^{\nu_2}(\mu)$. Then, we have

$$\mu \leq_c \mathcal{S}^{\nu_1}(\mu) \leq_c \mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_1}(\mu)) \leq \nu_2 \quad \text{and} \quad \mathcal{S}^{\nu_1}(\mu) \leq_c \mathcal{S}^{\nu_2}(\mu) \leq \nu_2.$$

These inequalities imply that

$$\mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_1}(\mu)) \in [\![\mu; \nu_2]\!] \quad \text{and} \quad \mathcal{S}^{\nu_2}(\mu) \in [\![\mathcal{S}^{\nu_1}(\mu); \nu_2]\!],$$

and now the minimality property of the shadow shows that

$$\mathcal{S}^{\nu_2}(\mu) \leq_c \mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_1}(\mu)) \quad \text{and} \quad \mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_1}(\mu)) \leq_c \mathcal{S}^{\nu_2}(\mathcal{S}^{\nu_2}(\mu)) = \mathcal{S}^{\nu_2}(\mu)$$

as desired. □

1.6.2 Construction of a Multistep Left-Monotone Transport

Our next goal is to define and construct a multistep left-monotone transport. The following concept will be crucial.

Definition 1.6.6. Let $\mu_0 \leq_{pc} \mu_1 \leq_c \cdots \leq_c \mu_n$. For $1 \leq t \leq n$, the *obstructed shadow of μ_0 in μ_t through $\mu_1, \ldots, \mu_{t-1}$* is iteratively defined by

$$\mathcal{S}^{\mu_1,\ldots,\mu_t}(\mu_0) := \mathcal{S}^{\mu_t}(\mathcal{S}^{\mu_1,\ldots,\mu_{t-1}}(\mu_0)).$$

The obstructed shadow is well-defined due to Lemma 1.6.5 (ii). An alternative definition is provided by the following characterization.

Lemma 1.6.7. *Let $\mu_0 \leq_{pc} \mu_1 \leq_c \cdots \leq_c \mu_n$ and $1 \leq t \leq n$. Then $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0)$ is the unique least element of the set*

$$[\![\mu_0;\mu_t]\!]^{\mu_1,\dots,\mu_{t-1}} := \{\theta_t \leq \mu_t : \exists \theta_s \leq \mu_s,\ 1 \leq s \leq t-1,\ \mu_0 \leq_c \theta_1 \leq_c \cdots \leq_c \theta_t\}$$

for the convex order; that is, $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0) \leq_c \theta$ for all elements θ.

Proof. For $t = 1$ this holds by the definition of the shadow in Lemma 1.6.1. For $t > 1$, we inductively assume that $\mathcal{S}^{\mu_1,\dots,\mu_{t-1}}(\mu_0)$ is the least element of $[\![\mu_0;\mu_{t-1}]\!]^{\mu_1,\dots,\mu_{t-2}}$. Consider an arbitrary element $\theta_t \in [\![\mu_0;\mu_t]\!]^{\mu_1,\dots,\mu_{t-1}}$ and fix some

$$\mu_0 \leq_c \theta_1 \leq_c \cdots \leq_c \theta_{t-1} \leq_c \theta_t \quad \text{with} \quad \theta_s \leq \mu_s, \quad 1 \leq s \leq t-1.$$

Then, $\theta_{t-1} \in [\![\mu_0;\mu_{t-1}]\!]^{\mu_1,\dots,\mu_{t-2}}$ and in particular $\mathcal{S}^{\mu_1,\dots,\mu_{t-1}}(\mu_0) \leq_c \theta_{t-1}$. Recall that $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0)$ is defined as the least element for $\leq_c$ of

$$\begin{aligned}[\![\mathcal{S}^{\mu_1,\dots,\mu_{t-1}}(\mu_0);\mu_t]\!] &= \{\theta \leq \mu_t : \mathcal{S}^{\mu_1,\dots,\mu_{t-1}}(\mu_0) \leq_c \theta\} \\ &\supseteq \{\theta \leq \mu_t : \theta_{t-1} \leq_c \theta\} \ni \theta_t.\end{aligned}$$

Hence, $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0) \leq_c \theta_t$, and as $\theta_t \in [\![\mu_0;\mu_t]\!]^{\mu_1,\dots,\mu_{t-1}}$ was arbitrary, this shows that $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0)$ is a least element of $[\![\mu_0;\mu_t]\!]^{\mu_1,\dots,\mu_{t-1}}$. The uniqueness of the least element follows from the general fact that $\theta_t^1 \leq_c \theta_t^2$ and $\theta_t^2 \leq_c \theta_t^1$ imply $\theta_t^1 = \theta_t^2$. $\square$

We can now state the main result of this section.

Theorem 1.6.8. *Let $\boldsymbol{\mu} = (\mu_0,\dots,\mu_n)$ be in convex order. Then there exists $P \in \mathcal{M}(\boldsymbol{\mu})$ such that the bivariate projections $P_{0t} := P \circ (X_0, X_t)^{-1}$ satisfy*

$$P_{0t}[(-\infty, x] \times A] = \mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{(-\infty,x]})(A) \quad \textit{for} \quad x \in \mathbb{R},\ A \in \mathfrak{B}(\mathbb{R}),$$

for all $1 \leq t \leq n$. Any such $P \in \mathcal{M}(\boldsymbol{\mu})$ is called a left-monotone transport.

We observe that an n-step left-monotone transport is defined purely in terms of its bivariate projections $P \circ (X_0, X_t)^{-1}$. In the one-step case, this completely determines the transport. For $n > 1$, we shall see that there can be multiple (and then infinitely many) left-monotone transports; in fact, they form a convex compact set. This will be discussed in more detail in Section 1.8, where it will also be shown that uniqueness does hold if μ_0 is atomless.

Proof of Theorem 1.6.8. Step 1. We first construct measures $\pi_t \in \Pi(\lambda, \mu_t)$, $0 \le t \le n$ such that

$$\pi_t|_{[0,\mu_0((-\infty,x])]\times\mathbb{R}} \circ X_1^{-1} = \mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{(-\infty,x]})$$

for all $x \in \mathbb{R}$, as well as measures $Q_t \in \Pi(\lambda, \mu_{t-1}, \mu_t)$, $1 \le t \le n$ such that

$$Q_t|_{[0,\mu_0((-\infty,x])]\times\mathbb{R}\times\mathbb{R}} \circ (X_1, X_2)^{-1} \in \mathcal{M}\big(\mathcal{S}^{\mu_1,\dots,\mu_{t-1}}(\mu_0|_{(-\infty,x]}), \mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{(-\infty,x]})\big) \tag{1.6.1}$$

for all $x \in \mathbb{R}$. Indeed, for $t = 0$, we take $\pi_0 \in \Pi(\lambda, \mu_0)$ to be the quantile[10] coupling. Then, applying Proposition 1.6.4 to π_0 yields the measure Q_1, and we can define $\pi_1 := Q_1 \circ (X_0, X_2)^{-1}$. Proceeding inductively, applying Proposition 1.6.4 to π_{t-1} yields Q_t which in turn allows us to define $\pi_t := Q_t \circ (X_0, X_2)^{-1}$.

Step 2. For $1 \le t \le n$, consider a disintegration $Q_t = \pi_{t-1} \otimes \kappa_t$ of Q_t. By (1.6.1), we may choose $\kappa_t(s, x_{t-1}, dx_t)$ to be a martingale kernel; that is,

$$\int x_t \, \kappa_t(s, x_{t-1}, dx_t) = x_{t-1}$$

holds for all $(s, x_{t-1}) \in \mathbb{R}^2$. We now define a measure $\pi \in \Pi(\lambda, \mu_0, \dots, \mu_n)$ on $\mathbb{R}^{n+2}$ via

$$\pi = \pi_0 \otimes \kappa_1 \otimes \cdots \otimes \kappa_n.$$

[10]The quantile coupling (or Fréchet–Hoeffding coupling) is given by the law of $(F_\lambda^{-1}, F_{\mu_0}^{-1})$ under λ, where $F_{\mu_0}^{-1}$ is the inverse c.d.f. of μ_0.

Then, π satisfies

$$\pi \circ (X_0, X_t)^{-1} = \pi_{t-1} \quad \text{and} \quad \pi \circ (X_0, X_t, X_{t+1})^{-1} = Q_t$$

for $1 \leq t \leq n$, and setting $P = \pi \circ (X_1, \ldots, X_{n+1})^{-1}$ yields the theorem. $\square$

The following result studies the bivariate projections P_{0t} of a left-monotone transport and shows in particular that P_{0t} may differ from the Left-Curtain coupling [16] in $\mathcal{M}(\mu_0, \mu_t)$.

Proposition 1.6.9. *Let $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ be in convex order and let $P \in \mathcal{M}(\boldsymbol{\mu})$ be a left-monotone transport. The following are equivalent:*

(i) The bivariate projection $P_{0t} = P \circ (X_0, X_t)^{-1} \in \mathcal{M}(\mu_0, \mu_t)$ is left-monotone for all $1 \leq t \leq n$.

(ii) The marginals $\boldsymbol{\mu}$ satisfy

$$\mathcal{S}^{\mu_1}(\mu_0|_{(-\infty,x]}) \leq_c \cdots \leq_c \mathcal{S}^{\mu_n}(\mu_0|_{(-\infty,x]}) \quad \textit{for all} \quad x \in \mathbb{R}. \tag{1.6.2}$$

Proof. Given $\mu \leq \mu_0$, an iterative application of Lemma 1.6.5 (ii) shows that the obstructed shadows coincide with the ordinary shadows, i.e. $\mathcal{S}^{\mu_1,\ldots,\mu_t}(\mu) = \mathcal{S}^{\mu_t}(\mu)$ for $1 \leq t \leq n$, if and only if $\mathcal{S}^{\mu_1}(\mu) \leq_c \cdots \leq_c \mathcal{S}^{\mu_n}(\mu)$. The proposition follows by applying this observation to $\mu = \mu_0|_{(-\infty,x]}$. $\square$

The following example illustrates the proposition and shows that (1.6.2) may indeed fail.

Example 1.6.10. Consider the marginals

$$\mu_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1, \quad \mu_1 = \frac{1}{2}\delta_{-2} + \frac{1}{2}\delta_2, \quad \mu_2 = \frac{1}{4}\delta_{-4} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_4.$$

Then the set $\mathcal{M}(\boldsymbol{\mu})$ consists of a single transport P; cf. the left panel of Figure 1.3. Thus, P is necessarily left-monotone. Similarly, $P_{01} = P \circ (X_0, X_1)^{-1}$ is the unique element of $\mathcal{M}(\mu_0, \mu_1)$. However, $P_{02} = P \circ (X_0, X_2)^{-1}$ is given by

$$\frac{3}{16}\delta_{(-1,-4)} + \frac{1}{4}\delta_{(-1,0)} + \frac{1}{16}\delta_{(-1,4)} + \frac{1}{16}\delta_{(1,-4)} + \frac{1}{4}\delta_{(1,0)} + \frac{3}{16}\delta_{(1,4)}$$

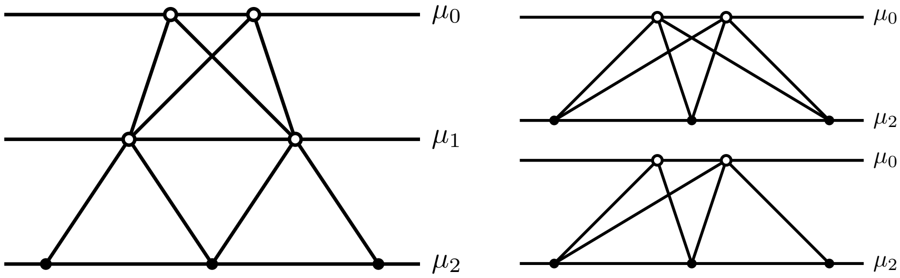


Figure 1.3: The left panel shows the support of the left-monotone transport P from Example 1.6.10. The right panel shows the support of P_{02} (top) and the support of the left-monotone transport in $\mathcal{M}(\mu_0, \mu_2)$ (bottom). The elements of the support are represented by the diagonal lines.

whereas the unique left-monotone transport in $\mathcal{M}(\mu_0, \mu_2)$ can be found to be

$$\frac{1}{8}\delta_{(-1,-4)} + \frac{3}{8}\delta_{(-1,0)} + \frac{1}{8}\delta_{(1,-4)} + \frac{1}{8}\delta_{(1,0)} + \frac{1}{4}\delta_{(1,4)}.$$

Therefore, there exists no transport $P \in \mathcal{M}(\boldsymbol{\mu})$ such that both P_{01} and P_{02} are left-monotone, and Proposition 1.6.9 shows that (1.6.2) fails.

Remark 1.6.11. Of course, all our results on left-monotone transports have "right-monotone" analogues, obtained by reversing the orientation on the real line (i.e. replacing $x \mapsto -x$ everywhere).

1.7 Geometry and Optimality Properties

In this section we introduce the optimality properties for transports and the geometric properties of their supports that were announced in the Introduction, and prove that they equivalently characterize left-monotone transports.

1.7.1 Geometry of Optimal Transports for Reward Functions of Spence–Mirrlees Type

The first goal is to show that optimal transports for specific reward functions are concentrated on sets $\Gamma \subseteq \mathbb{R}^{n+1}$ satisfying certain no-crossing conditions that we introduce next.

Given $1 \leq t \leq n$, we write

$$\Gamma^t = \{(x_0, \ldots, x_t) \in \mathbb{R}^{t+1} : (x_0, \ldots, x_n) \in \Gamma \text{ for some } (x_{t+1}, \ldots, x_n) \in \mathbb{R}^{n-t}\}$$

for the projection of Γ onto the first $t+1$ coordinates.

Definition 1.7.1. Let $\Gamma \subseteq \mathbb{R}^{n+1}$ and $1 \leq t \leq n$. Consider $\boldsymbol{x} = (x_0, \ldots, x_{t-1})$, $\boldsymbol{x}' = (x'_0, \ldots, x'_{t-1}) \in \mathbb{R}^t$ and $y^+, y^-, y' \in \mathbb{R}$ with $y^- < y^+$ such that $(\boldsymbol{x}, y^+), (\boldsymbol{x}, y^-), (\boldsymbol{x}', y') \in \Gamma^t$. Then, the projection

$$\Gamma^t \text{ is } \textit{left-monotone} \quad \text{if} \quad y' \notin (y^-, y^+) \text{ whenever } x_0 < x'_0.$$

The set Γ is left-monotone[11] if Γ^t is left-monotone for all $1 \leq t \leq n$.

We also need the following notion.

Definition 1.7.2. Let $\Gamma \subseteq \mathbb{R}^{n+1}$ and $1 \leq t \leq n$. The projection Γ^t is *nondegenerate* if for all $\boldsymbol{x} = (x_0, \ldots, x_{t-1}) \in \mathbb{R}^t$ and $y \in \mathbb{R}$ such that $(\boldsymbol{x}, y) \in \Gamma^t$, the following hold:

(i) if $y > x_{t-1}$, there exists $y' < x_{t-1}$ such that $(\boldsymbol{x}, y') \in \Gamma^t$;

(ii) if $y < x_{t-1}$, there exists $y' > x_{t-1}$ such that $(\boldsymbol{x}, y') \in \Gamma^t$.

The set Γ is called nondegenerate[12] if Γ^t is nondegenerate for all $1 \leq t \leq n$.

Broadly speaking, this definition says that for any path to the right in Γ there exists a path to the left, and vice versa. For a set supporting a martingale, nondegeneracy is not a restriction, in the following sense.

Remark 1.7.3. Let $\boldsymbol{\mu}$ be in convex order, $\mathcal{V}$ its effective domain and $\Gamma \subseteq \mathcal{V}$.

(i) There exists a nondegenerate, universally measurable set $\Gamma' \subseteq \Gamma$ such that $P(\Gamma') = 1$ for all $P \in \mathcal{M}(\boldsymbol{\mu})$ with $P(\Gamma) = 1$.

(ii) Fix $P \in \mathcal{M}(\boldsymbol{\mu})$ with $P(\Gamma) = 1$. There exists a nondegenerate, Borel-measurable set $\Gamma'_P \subseteq \Gamma$ such that $P(\Gamma'_P) = 1$.

[11] This terminology for Γ is abusive since $\Gamma = \Gamma^n$ is in fact a projection itself—it will be clear from the context what is meant.

[12] Footnote 11 applies here as well.

Proof. Let N_t be the set of all $\boldsymbol{x} \in \Gamma^t$ such that (i) or (ii) of Definition 1.7.2 fail. If P is a martingale with $P(\Gamma) = 1$, we see that $N_t \times \mathbb{R}^{n-t+1}$ is P-null. Moreover, N_t is universally measurable (as the projection of a Borel set) and we can set

$$\Gamma' := \Gamma \setminus \bigcup_{t=1}^{n} (N_t \times \mathbb{R}^{n-t+1})$$

to prove (i). Turning to (ii), universal measurability implies that there exists a Borel set $N_t' \supseteq N_t$ such that $N_t' \setminus N_t$ is P_{t-1}-null, where $P_{t-1} = P \circ (X_0, \dots, X_{t-1})^{-1}$. We can then set $\Gamma_P' := \Gamma \setminus \cup_{t=1}^{n} (N_t' \times \mathbb{R}^{n-t+1})$. $\square$

Next, we introduce a notion of competitors along the lines of [16, Definition 1.10].

Definition 1.7.4. Let π be a finite measure on $\mathbb{R}^{t+1}$ whose marginals have finite first moments and consider a disintegration $\pi = \pi_t \otimes \kappa$, where π_t is the projection of π onto the first t coordinates. A measure $\pi' = \pi_t \otimes \kappa'$ is a *t-competitor* of π if it has the same last marginal and

$$\mathrm{bary}(\kappa(\boldsymbol{x}, \cdot)) = \mathrm{bary}(\kappa'(\boldsymbol{x}, \cdot)) \quad \text{for } \pi_t\text{-a.e.} \quad \boldsymbol{x} = (x_0, \dots, x_{t-1}).$$

Using these definitions, we now formulate a variant of the monotonicity principle stated in Theorem 1.5.4 (i) that will be convenient to infer the geometry of Γ.

Lemma 1.7.5. *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order, $1 \le t \le n$ and let $\bar{f} : \mathbb{R}^{t+1} \to [0, \infty)$ be Borel. Consider $f(X_0, \dots, X_n) := \bar{f}(X_0, \dots, X_t)$ and suppose that $\mathbf{I}_{\boldsymbol{\mu}}(f) < \infty$. Let $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f)$ be an optimizer for $\mathbf{I}_{\boldsymbol{\mu}}(f)$ with the property that $\phi_s \equiv H_s \equiv 0$ for $s = t+1, \dots, n$ and define the set*

$$\Gamma := \left\{ \sum_{t=0}^{n} \phi_t(X_t) + (H \cdot X)_n = f \right\} \cap \mathcal{V}.$$

Let π be a finitely supported probability on $\mathbb{R}^{t+1}$ which is concentrated on Γ^t. Then $\pi(\bar{f}) \ge \pi'(\bar{f})$ for any t-competitor π' of π that is concentrated on $\mathcal{V}^t$.

Proof. Recall that the projections π_t and π'_t onto the first t coordinates coincide. Thus,

$$\begin{aligned}
\pi[H_t \cdot (X_t - X_{t-1})] &= \int H_t \cdot (\mathrm{bary}(\kappa(X_0, \dots, X_{t-1}, \cdot) - X_{t-1}) d\pi_t \\
&= \int H_t \cdot (\mathrm{bary}(\kappa'(X_0, \dots, X_{t-1}, \cdot) - X_{t-1}) d\pi'_t \\
&= \pi'[H_t \cdot (X_t - X_{t-1})].
\end{aligned}$$

Using also that the last marginals coincide, we deduce that

$$\pi[\bar{f}] = \pi\left[\sum_{s=0}^{t} \phi_s(X_s) + (H \cdot X)_t\right] = \pi'\left[\sum_{s=0}^{t} \phi_s(X_s) + (H \cdot X)_t\right] \geq \pi'[\bar{f}].$$

□

Next, we formulate an intermediate result relating optimality for Spence–Mirrlees reward functions to left-monotonicity of the support.

Lemma 1.7.6. *Let $1 \leq t \leq n$ and let $\Gamma \subseteq \mathcal{V}$ be a subset such that Γ^t is nondegenerate. Moreover, let $f : \mathbb{R}^{t+1} \to \mathbb{R}$ be of the form $f(X_0, \dots, X_t) = \bar{f}(X_0, X_t)$ for a second-order Spence–Mirrlees function $\bar{f}$. Assume that for any finitely supported probability π that is concentrated on Γ^t and any t-competitor π' of π that is concentrated on $\mathcal{V}^t$, we have $\pi(f) \geq \pi'(f)$. Then, the projection Γ^t is left-monotone.*

Proof. Consider $(\boldsymbol{x}, y_1), (\boldsymbol{x}, y_2), (\boldsymbol{x}', y') \in \Gamma^t$ satisfying $x_0 < x'_0$ and suppose for contradiction that $y_1 < y' < y_2$. We define $\lambda = \frac{y_2 - y'}{y_2 - y_1}$ and

$$\begin{aligned}
\pi &= \frac{\lambda}{2}\delta_{(\boldsymbol{x}, y_1)} + \frac{1-\lambda}{2}\delta_{(\boldsymbol{x}, y_2)} + \frac{1}{2}\delta_{(\boldsymbol{x}', y')} \\
\pi' &= \frac{\lambda}{2}\delta_{(\boldsymbol{x}', y_1)} + \frac{1-\lambda}{2}\delta_{(\boldsymbol{x}', y_2)} + \frac{1}{2}\delta_{(\boldsymbol{x}, y')}.
\end{aligned}$$

Then π and π' have the same projection $\pi_t = \pi'_t$ on the first t marginals and their last marginals also coincide. Moreover, disintegrating $\pi = \pi_t \otimes \kappa$ and $\pi' = \pi_t \otimes \kappa'$, the measures $\kappa(\boldsymbol{x}), \kappa(\boldsymbol{x}'), \kappa'(\boldsymbol{x}), \kappa(\boldsymbol{x}')$ all have barycenter y'. Therefore, π and π' are t-competitors. We must also have that π' is concentrated on $\mathcal{V}^t$, by the shape of $\mathcal{V}$. Now our assumption implies that $\pi(f) \geq \pi'(f)$, but the second-order Spence–Mirrlees property of $\bar{f}$ implies that $\pi(f) < \pi'(f)$. □

1.7.2 Geometry of Left-Monotone Transports

Next, we establish that transports with left-monotone support are indeed left-monotone in the sense of Theorem 1.6.8.

Theorem 1.7.7. *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order and let $P \in \mathcal{M}(\boldsymbol{\mu})$ be concentrated on a nondegenerate, left-monotone set $\Gamma \subseteq \mathbb{R}^{n+1}$. Then P is left-monotone.*

Before stating the proof of the theorem, we record two auxiliary results about measures on the real line. The first one is a direct consequence of Proposition 1.2.1.

Lemma 1.7.8. *Let $a < b$ and $\mu \leq_c \nu$. If ν is concentrated on $(-\infty, a]$, then so is μ, and moreover $\nu(\{a\}) \geq \mu(\{a\})$. The analogue holds for $[b, \infty)$.*

The second result is [16, Lemma 5.2].

Lemma 1.7.9. *Let σ be a nontrivial signed measure on $\mathbb{R}$ with $\sigma(\mathbb{R}) = 0$ and let $\sigma = \sigma^+ - \sigma^-$ be its Hahn decomposition. There exist $a \in \operatorname{supp}(\sigma^+)$ and $b > a$ such that $\int (b - y)^+ \mathbf{1}_{[a,\infty)} \, d\sigma(y) > 0$.*

We can now give the proof of the theorem; it is inspired by [16, Theorem 5.3] which corresponds to the case $n = 1$.

Proof of Theorem 1.7.7. Since the case $n = 1$ is covered by Proposition 1.6.3, we may assume that the theorem has been proved for transports with $n - 1$ steps and focus on the induction argument.

For every $x \in \mathbb{R}$ we denote by μ_x^t the marginal $(P|_{(-\infty,x]\times\mathbb{R}^n}) \circ X_t^{-1}$. In particular, we then have $\mu_x^0 = \mu_0|_{(-\infty,x]}$ and μ_x^t is the image of μ_x^0 under P after t steps. For the sake of brevity, we also set $\nu_x^t := \mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_x^0)$. By definition, P is left-monotone if $\mu_x^t = \nu_x^t$ for all $x \in \mathbb{R}$ and $t \leq n$, and by the induction hypothesis, we may assume that this holds for $t \leq n - 1$.

We argue by contradiction and assume that there exists $x \in \mathbb{R}$ such that $\mu_x^n \neq \nu_x^n$. Then, the signed measure

$$\sigma := \nu_x^n - \mu_x^n$$

is nontrivial and we can find $a < b$ with $a \in \operatorname{supp}(\sigma^+)$ as in Lemma 1.7.9. Observe that $\sigma^+ \leq \mu_n - \mu_x^n$ where $\mu_n - \mu_x^n$ is the image of $\mu_n|_{(x,\infty)}$ under P. Hence, $a \in \operatorname{supp}(\mu_n - \mu_x^n)$

and as P is concentrated on Γ, we conclude that there exists a sequence of points

$$\boldsymbol{x}^m = (x_0^m, \dots, x_n^m) \in \Gamma \quad \text{with } x < x_0^m \text{ and } x_n^m \to a. \tag{1.7.1}$$

Moreover, by the characterization of the obstructed shadow in Lemma 1.6.7, we must have

$$\nu_x^n \leq_c \mu_x^n$$

as $\mu_x^n \in [\![\mu_x^0; \mu_n]\!]^{\mu_1,\dots,\mu_{n-1}}$ due to the fact that μ_x^n is the image of μ_x^0 under a martingale transport.

Step 1. We claim that for all $\boldsymbol{x} = (x_0, \dots, x_{n-1})$ with $x_0 \leq x$ and $x_{n-1} \leq a$, it holds that

$$\Gamma_{\boldsymbol{x}} \cap (a, \infty) = \emptyset,$$

where $\Gamma_{\boldsymbol{x}} = \{y \in \mathbb{R} : (\boldsymbol{x}, y) \in \Gamma\}$ is the section of Γ at $\boldsymbol{x}$. By way of contradiction, assume that for some $\boldsymbol{x}$ with $x_0 \leq x$ and $x_{n-1} \leq a$ we have $\Gamma_{\boldsymbol{x}} \cap (a, \infty) \neq \emptyset$, then in particular $\Gamma_{\boldsymbol{x}} \cap (x_{n-1}, \infty) \neq \emptyset$. In view of the nondegeneracy of Γ, we conclude that $\Gamma_{\boldsymbol{x}} \cap (-\infty, x_{n-1}) \neq \emptyset$ and hence that $\Gamma_{\boldsymbol{x}} \cap (-\infty, a) \neq \emptyset$. This yields a contradiction to the left-monotonicity of Γ by using $\boldsymbol{x}^m$ from (1.7.1) for $\boldsymbol{x}'$ in Definition 1.7.1 for large enough m, and the proof of the claim is complete.

Step 2. Similarly, we can show that for all $\boldsymbol{x} = (x_0, \dots, x_{n-1})$ with $x_0 \leq x$ and $x_{n-1} \geq a$,

$$\Gamma_{\boldsymbol{x}} \cap (-\infty, a) = \emptyset.$$

Step 3. Next, we consider the marginals

$$\mu_{x,a}^t := \left(P|_{(-\infty,x]\times\mathbb{R}^{n-2}\times(-\infty,a]\times\mathbb{R}}\right) \circ X_t^{-1}.$$

Then, in particular, $\mu_{x,a}^{n-1} = \mu_x^{n-1}|_{(-\infty,a]}$ and $\mu_{x,a}^n$ is the image of $\mu_{x,a}^{n-1}$ under the last step of P. Step 1 of the proof thus implies that $\mu_{x,a}^n$ is concentrated on $(-\infty, a]$. We also write

$$\nu_{x,a}^n := \mathcal{S}^{\mu_n}(\mu_x^{n-1}|_{(-\infty,a]}).$$

We have $\mu_{x,a}^{n-1} \leq_c \mu_{x,a}^n$ as $\mathcal{M}(\mu_{x,a}^{n-1}, \mu_{x,a}^n) \neq \emptyset$, and $\mu_{x,a}^n \leq \mu_x^n \leq \mu_n$. Therefore,

$$\nu_{x,a}^n \leq_c \mu_{x,a}^n \tag{1.7.2}$$

by the minimality of the shadow. Next, we show that

$$\nu_x^n - \nu_{x,a}^n \leq_c \mu_x^n - \mu_{x,a}^n. \tag{1.7.3}$$

Observe that $\mu_x^n - \mu_{x,a}^n$ is the image of $\mu_x^{n-1}|_{(a,\infty)}$ under P and therefore concentrated on $[a, \infty)$ by Step 2. Using this observation, that $\mu_{x,a}^n$ is concentrated on $(-\infty, a]$ as mentioned above, and the fact that $\nu_{x,a}^n(\{a\}) \leq \mu_{x,a}^n(\{a\})$ as a consequence of (1.7.2) and Lemma 1.7.8, we have

$$\mu_x^n - \mu_{x,a}^n = (\mu_x^n - \mu_{x,a}^n)|_{[a,\infty)} \leq (\mu_n - \mu_{x,a}^n)|_{[a,\infty)} \leq (\mu_n - \nu_{x,a}^n)|_{[a,\infty)} \leq \mu_n - \nu_{x,a}^n.$$

We also have $\mu_x^{n-1}|_{(a,\infty)} \leq_c \mu_x^n - \mu_{x,a}^n$ since the latter measure is the image of the former under P. Together with the preceding display, we have established that

$$\mu_x^n - \mu_{x,a}^n \in \left[\!\left[\mu_x^{n-1}|_{(a,\infty)}; \mu_n - \nu_{x,a}^n\right]\!\right].$$

On the other hand,

$$\nu_x^n - \nu_{x,a}^n = \mathcal{S}^{\mu_n - \nu_{x,a}^n}(\mu_x^{n-1}|_{(a,\infty)})$$

from the additivity property of the shadow in Lemma 1.6.5 (i), and therefore (1.7.3) follows by the minimality of the shadow.

Step 4. Recall from Step 3 that $\mu_{x,a}^n$ is concentrated on $(-\infty, a]$ and that $\mu_x^n - \mu_{x,a}^n$ is concentrated on $[a, \infty)$. Therefore, $\nu_{x,a}^n$ is concentrated on $(-\infty, a]$ and $\nu_x^n - \nu_{x,a}^n$ is concentrated on $[a, \infty)$, by Lemma 1.7.8. Moreover, we have $\nu_{x,a}^n(\{a\}) \leq \mu_{x,a}^n(\{a\})$ by the same lemma, and finally, the function $y \mapsto (b - y)^+ \mathbf{1}_{[a,\infty)}(y)$ is convex on $[a, \infty)$ as $a < b$.

Using these facts and (1.7.3),

$$\begin{aligned}\int (b-y)^+ \mathbf{1}_{[a,\infty)}(y)\nu_x^n(dy) \\ &= \int (b-y)^+ \mathbf{1}_{[a,\infty)}(y)(\nu_x^n - \nu_{x,a}^n)(dy) + (b-a)\nu_{x,a}^n(\{a\}) \\ &\leq \int (b-y)^+ \mathbf{1}_{[a,\infty)}(y)(\mu_x^n - \mu_{x,a}^n)(dy) + (b-a)\mu_{x,a}^n(\{a\}) \\ &= \int (b-y)^+ \mathbf{1}_{[a,\infty)}(y)\mu_x^n(dy).\end{aligned}$$

This contradicts the choice of a and b, cf. Lemma 1.7.9, and thus completes the proof. $\square$

1.7.3 Optimality Properties

In this section we relate left-monotone transports and left-monotone sets to the optimal transport problem for Spence–Mirrlees functions.

Theorem 1.7.10. *For $1 \leq t \leq n$, let $f_t : \mathbb{R}^2 \to \mathbb{R}$ be second-order Spence–Mirrlees functions such that $|f_t(x,y)| \leq a_0(x) + a_t(y)$ for some $a_0 \in L^1(\mu_0)$ and $a_t \in L^1(\mu_t)$. There exists a universally measurable, nondegenerate, left-monotone set $\Gamma' \subseteq \mathbb{R}^{n+1}$ such that any simultaneous optimizer $P \in \mathcal{M}(\boldsymbol{\mu})$ for $\mathbf{S}_{\boldsymbol{\mu}}(f_t(X_0, X_t))$, $1 \leq t \leq n$ is concentrated on Γ'. In particular, any such P is left-monotone.*

Proof. The last assertion follows by an application of Theorem 1.7.7, so we may focus on finding Γ'. For each $1 \leq t \leq n$, we use Theorem 1.5.2 and Remark 1.5.3 to find a dual optimizer $(\boldsymbol{\phi}, H) \in \mathcal{D}_{\boldsymbol{\mu}}(f_t)$ for $\mathbf{I}_{\boldsymbol{\mu}}(f_t(X_0, X_t))$ and define the Borel set

$$\Gamma(t) := \left\{ \sum_{s=0}^{n} \phi_s(X_s) + (H \cdot X)_n = f_t \right\} \cap \mathcal{V}.$$

Here, we may choose a dual optimizer such that $\phi_s \equiv H_s \equiv 0$ for $s = t+1, \dots, n$. (This can be seen by applying Theorem 1.5.2 to the transport problem involving only the marginals $(\mu_0, \dots, \mu_t)$ and taking the corresponding dual optimizer.) Theorem 1.5.4 shows that any simultaneous optimizer $P \in \mathcal{M}(\boldsymbol{\mu})$ is concentrated on $\Gamma(t)$ for all t, and hence also on the Borel set

$$\Gamma := \bigcap_{t=1}^{n} \Gamma(t).$$

Using Remark 1.7.3 (i), we find a universally measurable, nondegenerate subset $\Gamma' \subseteq \Gamma$ with the same property. Since the projection $(\Gamma')^t$ is contained in the projection $(\Gamma(t))^t$, Lemma 1.7.5 and Lemma 1.7.6 yield that $(\Gamma')^t$ is left-monotone for all t; that is, Γ' is left-monotone. $\square$

Remark 1.7.11. In Theorem 1.7.10, if we only wish to find a nondegenerate, left-monotone set $\Gamma'_P \subseteq \mathbb{R}^{n+1}$ such that a given simultaneous optimizer $P \in \mathcal{M}(\boldsymbol{\mu})$ is concentrated on Γ'_P, then we may choose Γ'_P to be Borel instead of universally measurable. This follows by replacing the application of Remark 1.7.3 (i) by Remark 1.7.3 (ii) in the proof.

The following is a converse to Theorem 1.7.10.

Theorem 1.7.12. *Given $1 \leq t \leq n$, let $f \in C^{1,2}(\mathbb{R}^2)$ be such that $f_{xyy} \geq 0$ and suppose that the following integrability condition holds:*

$$\begin{cases} f(X_0, X_t), \quad f(0, X_t), \quad f(X_0, 0), \quad \bar{h}(X_0)X_0, \quad \bar{h}(X_0)X_t \\ \text{are } P\text{-integrable for all } \; P \in \mathcal{M}(\boldsymbol{\mu}), \end{cases} \tag{1.7.4}$$

where $\bar{h}(x) := \partial_y|_{y=0}[f(x, y) - f(0, y)]$. Then every left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ is an optimizer for $\mathbf{S}_{\boldsymbol{\mu}}(f)$.

The integrability condition clearly holds when f is Lipschitz continuous; in particular, a smooth second-order Spence–Mirrlees function (as defined in the Introduction) satisfies the assumptions of the theorem for any $\boldsymbol{\mu}$.

The proof will be given by an approximation based on the following building blocks for Spence–Mirrlees functions; the construction is novel and may be of independent interest.

Lemma 1.7.13. *Let $1 \leq t \leq n$ and let $f(X_0, \dots, X_n) := \mathbf{1}_{(-\infty, a]}(X_0)\varphi(X_t)$ for a concave function φ and $a \in \mathbb{R}$. Then every left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ is an optimizer for $\mathbf{S}_{\boldsymbol{\mu}}(f)$.*

Proof. In view of Lemma 1.6.7, this follows directly by applying the defining shadow property from Theorem 1.6.8 with $x = a$. $\square$

The integrability condition (1.7.4) implies that setting

$$g(x, y) := f(x, 0) + f(0, y) - f(0, 0) + \bar{h}(x)y,$$

the three terms constituting

$$g(X_0, X_t) = [f(X_0, 0) + \bar{h}(X_0)X_0] + [f(0, X_t) - f(0, 0)] + [\bar{h}(X_0)(X_t - X_0)]$$

are P-integrable and $P[g(X_0, X_t)]$ is constant over $P \in \mathcal{M}(\boldsymbol{\mu})$. By replacing f with $f - g$, we may thus assume without loss of generality that

$$f(x, 0) = f(0, y) = f_y(x, 0) = 0 \quad \text{for all} \quad (x, y) \in \mathbb{R}^2. \tag{1.7.5}$$

After this normalization, integration by parts yields the representation

$$f(x, y) = \int_0^y \int_0^x (y - t) f_{xyy}(s, t)\, ds\, dt. \tag{1.7.6}$$

Lemma 1.7.14. *Theorem 1.7.12 holds under the following additional condition: there exists a constant $c > 0$ such that*

$$x \mapsto f(x, y) \textit{ is constant on } \{x > c\} \textit{ and on } \{x < -c\},$$
$$y \mapsto f(x, y) \textit{ is affine on } \{y > c\} \textit{ and on } \{y < -c\}.$$

Proof. Integration by parts implies that for all $(x, y) \in \mathbb{R}^2$, we have the representation

$$\begin{aligned} f(x, y) = &- \int_{-c}^{c} \int_{-c}^{c} \mathbf{1}_{(-\infty, s]}(x)(y - t)^+ f_{xyy}(s, t)\, ds\, dt \\ &+ [f(x, -c) - (-c) f_y(x, -c)] \\ &+ [f(c, y) - f(c, -c) - f_y(c, -c)(y - (-c))] \\ &+ f_y(x, -c) y. \end{aligned}$$

The last three terms are of the form $g(x, y) = \tilde{\phi}(x) + \tilde{\psi}(y) + \tilde{h}(x)y$ and of linear growth due to the additional condition. Hence, as above, $P'[g(X_0, X_t)] = C$ is constant for $P' \in \mathcal{M}(\boldsymbol{\mu})$. If

$P \in \mathcal{M}(\boldsymbol{\mu})$ is left-monotone and $P' \in \mathcal{M}(\boldsymbol{\mu})$ is arbitrary, Fubini's theorem and Lemma 1.7.13 yield that

$$
\begin{aligned}
P[f] &= -\int_{-c}^{c}\int_{-c}^{c} P[\mathbf{1}_{(-\infty,s]}(x)(y-t)^{+}]f_{xyy}(s,t)\,ds\,dt + C \\
&\geq -\int_{-c}^{c}\int_{-c}^{c} P'[\mathbf{1}_{(-\infty,s]}(x)(y-t)^{+}]f_{xyy}(s,t)\,ds\,dt + C \\
&= P'[f],
\end{aligned}
$$

where P, P' are understood to integrate with respect to (x,y) and the application of Fubini's theorem is justified by the nonnegativity of the integrand. $\square$

Proof of Theorem 1.7.12. Let f be as in the theorem. We shall construct functions f^m, $m \geq 1$ satisfying the assumption of Lemma 1.7.14 as well as $P[f^m] \to P[f]$ for all $P \in \mathcal{M}(\boldsymbol{\mu})$. Once this is achieved, the theorem follows from the lemma.

Indeed, we may assume that f is normalized as in (1.7.5). Let $m \geq 1$ and let $\rho_m : \mathbb{R} \to [0,1]$ be a smooth function such that $\rho_m = 1$ on $[-m,m]$ and $\rho_m = 0$ on $[-m-1,m+1]^c$. In view of (1.7.6), we define f^m by

$$
f^m(x,y) = \int_0^y \int_0^x (y-t) f_{xyy}(s,t)\rho_m(s)\rho_m(t)\,ds\,dt.
$$

It then follows that f^m satisfies the assumptions of Lemma 1.7.14 with the constant $c = m+1$. Moreover, we have

$$
0 \leq f^m(x,y) \leq f^{m+1}(x,y) \leq f(x,y) \quad \text{for} \quad x \geq 0
$$

and the opposite inequalities for $x \leq 0$, as well as $f^m(x,y) \to f(x,y)$ for all (x,y).

Let $P \in \mathcal{M}(\boldsymbol{\mu})$. Since f is P-integrable, applying monotone convergence separately on $\{x \geq 0\}$ and $\{x \leq 0\}$ yields that $P[f^m] \to P[f]$, and the proof is complete. $\square$

Remark 1.7.15. The function

$$
\bar{f}(x,y) := \tanh(x)\sqrt{1+y^2}
$$

satisfies the conditions of Theorem 1.7.12 for all marginals $\boldsymbol{\mu}$ in convex order, since the latter are assumed to have a finite first moment.

We can now collect the preceding results to obtain, in particular, the equivalences stated in Theorem 1.1.1.

Theorem 1.7.16. *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order. There exists a left-monotone, nondegenerate, universally measurable set $\Gamma \subseteq \mathbb{R}^{n+1}$ such that for any $P \in \mathcal{M}(\boldsymbol{\mu})$, the following are equivalent:*

(i) P is an optimizer for $\mathbf{S}_{\boldsymbol{\mu}}(f(X_0, X_t))$ whenever f is a smooth second-order Spence–Mirrlees function and $1 \leq t \leq n$,

(ii) P is concentrated on Γ,

(ii') P is concentrated on a left-monotone set,

(iii) P is left-monotone; i.e. P_{0t} transports $\mu_0|_{(-\infty,a]}$ to $\mathcal{S}^{\mu_1,\dots,\mu_t}(\mu_0|_{(-\infty,a]})$ for all $1 \leq t \leq n$ and $a \in \mathbb{R}$.

Moreover, there exists $P \in \mathcal{M}(\boldsymbol{\mu})$ satisfying (i)–(iii).

Proof. Let Γ be the set provided by Theorem 1.7.10 for the function $f_t = \bar{f}$ of Remark 1.7.15. Given $P \in \mathcal{M}(\boldsymbol{\mu})$, Theorem 1.7.10 shows that (i) implies (ii) which trivially implies (ii'). Theorem 1.7.7 and Remark 1.7.3 show that (ii') implies (iii), and Theorem 1.7.12 shows that (iii) implies (i). Finally, the existence of a left-monotone transport was stated in Theorem 1.6.8. ☐

We conclude this section with an example showing that left-monotone transports are not Markovian in general, even if they are unique and (1.6.2) holds for $\boldsymbol{\mu}$.

Example 1.7.17. Consider the marginals

$$\mu_0 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \quad \mu_1 = \frac{3}{4}\delta_0 + \frac{1}{4}\delta_2, \quad \mu_2 = \frac{1}{8}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1}{8}\delta_1 + \frac{1}{4}\delta_2.$$

The transport $P \in \mathcal{M}(\boldsymbol{\mu})$ given by

$$P = \frac{1}{2}\delta_{(0,0,0)} + \frac{1}{8}\delta_{(1,0,-1)} + \frac{1}{8}\delta_{(1,0,1)} + \frac{1}{4}\delta_{(1,2,2)}$$

is left-monotone because its support is left-monotone (Figure 1.4), and it is clearly not Markovian. On the other hand, it is not hard to see that this is the only way to build a left-monotone transport in $\mathcal{M}(\boldsymbol{\mu})$.

1.8 Uniqueness of Left-Monotone Transports

In this section we consider the (non-)uniqueness of left-monotone transports. It turns out the presence of atoms in μ_0 is important in this respect—let us start with the following simple observation.

Remark 1.8.1. Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order. If μ_0 is a Dirac mass, then every $P \in \mathcal{M}(\boldsymbol{\mu})$ is left-monotone. Indeed, $\mathcal{M}(\mu_0, \mu_t)$ is a singleton for every $1 \leq t \leq n$, hence P_{0t} must be the (one-step) left-monotone transport.

Exploiting this observation, the following shows that left-monotone transports need not be unique when $n \geq 2$.

Example 1.8.2. Let $\mu_0 = \delta_0$, $\mu_1 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, $\mu_2 = \frac{3}{8}\delta_{-2} + \frac{1}{4}\delta_0 + \frac{3}{8}\delta_2$. By the remark, any element in $\mathcal{M}(\boldsymbol{\mu})$ is left-monotone. Moreover, $\mathcal{M}(\boldsymbol{\mu})$ is a continuum since $\mathcal{M}(\mu_1, \mu_2)$ contains the convex hull of the two measures

$$
\begin{aligned}
P_l &= \frac{1}{4}\delta_{(-1,-2)} + \frac{1}{4}\delta_{(-1,0)} + \frac{1}{8}\delta_{(1,-2)} + \frac{3}{8}\delta_{(1,2)}, \\
P_r &= \frac{3}{8}\delta_{(-1,-2)} + \frac{1}{8}\delta_{(-1,2)} + \frac{1}{4}\delta_{(1,0)} + \frac{1}{4}\delta_{(1,2)}.
\end{aligned}
$$

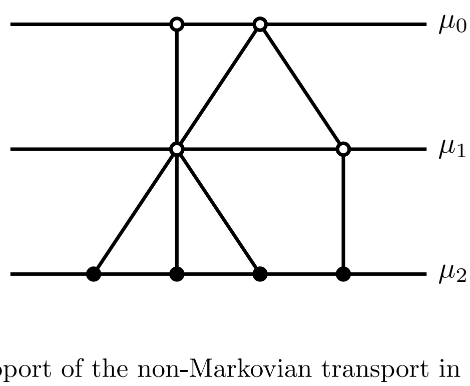


Figure 1.4: Support of the non-Markovian transport in Example 1.7.17.

The corresponding supports are depicted in Figure 1.5.

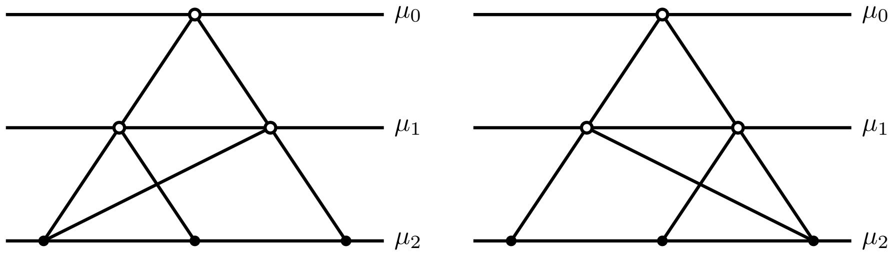


Figure 1.5: Supports of two left-monotone transports for the same marginals.

The example illustrates that non-uniqueness can typically be expected when μ_0 has atoms. On the other hand, we have the following uniqueness result.

Theorem 1.8.3. *Let $\boldsymbol{\mu} = (\mu_0, \dots, \mu_n)$ be in convex order. If μ_0 is atomless, there exists a unique left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$.*

The remainder of this section is devoted to the proof. Let us call a kernel $\kappa(x, dy)$ *binomial* if for all $x \in \mathbb{R}$, the measure $\kappa(x, dy)$ consists of (at most) two point masses. A martingale transport will be called binomial if it can be disintegrated using only binomial kernels. We shall show that when μ_0 is atomless, any left-monotone transport is a binomial martingale, and then conclude the uniqueness via a convexity argument.

The first step is the following set-theoretic result.

Lemma 1.8.4. *Let $k \geq 1$ be an integer and $\Gamma \subseteq \mathbb{R}^{t+1}$. For $\boldsymbol{x} \in \mathbb{R}^t$, we denote by $\Gamma_{\boldsymbol{x}} := \{y \in \mathbb{R} : (\boldsymbol{x}, y) \in \Gamma\}$ the section at $\boldsymbol{x}$. If the set*

$$\{\boldsymbol{x} \in \mathbb{R}^t : |\Gamma_{\boldsymbol{x}}| \geq k\}$$

is uncountable, then it has an accumulation point. More precisely, there are $\boldsymbol{x} = (x_0, \dots, x_t) \in \mathbb{R}^t$ and $y_1 < \cdots < y_k$ in $\Gamma_{\boldsymbol{x}}$ such that for all $\epsilon > 0$ there exist $\boldsymbol{x}' = (x'_0, \dots, x'_t) \in \mathbb{R}^t$ and $y'_1 < \cdots < y'_k$ in $\Gamma_{\boldsymbol{x}'}$ satisfying

(i) $\|\boldsymbol{x} - \boldsymbol{x}'\| < \epsilon$,

(ii) $x_0 < x'_0$,

(iii) $\max_{i=1,\ldots,k} |y_i - y_i'| < \epsilon$.

Proof. The proof is similar to the one of [16, Lemma 3.2] and therefore omitted. □

The following statement on the binomial structure generalizes a result of [16] for the one-step case and is of independent interest.

Proposition 1.8.5. *Let $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ be in convex order and let μ_0 be atomless. There exists a universally measurable set $\Gamma \subseteq \mathbb{R}^{n+1}$ such that every left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ is concentrated on Γ and such that for all $1 \leq t \leq n$ and $\boldsymbol{x} \in \mathbb{R}^t$,*

$$|\{y \in \mathbb{R} : (X_0, \ldots, X_t)^{-1}(\boldsymbol{x}, y) \cap \Gamma \neq \emptyset\}| \leq 2. \tag{1.8.1}$$

In particular, every left-monotone transport $P \in \mathcal{M}(\boldsymbol{\mu})$ is a binomial martingale.

Proof. Let Γ be as in Theorem 1.7.16; then every left-monotone $P \in \mathcal{M}(\boldsymbol{\mu})$ is concentrated on Γ. Let A_t be the set of all $\boldsymbol{x} \in \mathbb{R}^t$ such that (1.8.1) fails. Suppose that A_t is uncountable; then Lemma 1.8.4 yields points $\boldsymbol{x}, \boldsymbol{x}'$ such that for some $y_1, y_2 \in \Gamma^t_{\boldsymbol{x}}$ and $y \in \Gamma^t_{\boldsymbol{x}'}$ we have $y_1 < y < y_2$. This contradicts the left-monotonicity of Γ (Definition 1.7.1), thus A_t must be countable. Hence, $(X_0, \ldots, X_{t-1})^{-1}(A_t)$ is Borel and P-null for all $P \in \mathcal{M}(\boldsymbol{\mu})$, as μ_0 is atomless. The set $\Gamma' = \Gamma \setminus \cup_{t=1}^n (X_0, \ldots, X_{t-1})^{-1}(A_t)$ then has the required properties. □

Proof of Theorem 1.8.3. We will prove this result using induction on n. For $n = 1$ the result holds by Proposition 1.6.3, with or without atoms. To show the induction step, let P' be the unique left-monotone transport in $\mathcal{M}(\mu_0, \ldots, \mu_{n-1})$ and let $P_1 = P' \otimes \kappa_1$ and $P_2 = P' \otimes \kappa_2$ be disintegrations of two n-step left-monotone transports. Then,

$$\frac{P_1 + P_2}{2} = P' \otimes \frac{\kappa_1 + \kappa_2}{2}$$

is again left-monotone, and Proposition 1.8.5 yields that $(\kappa_1 + \kappa_2)/2$ must be a binomial kernel P'-a.s. Using also the martingale property of κ_1 and κ_2, this can only be true if $\kappa_1 = \kappa_2$ holds P'-a.s., and therefore $P_1 = P_2$. □

1.9 Free Intermediate Marginals

In this section we discuss a variant of our transport problem where the intermediate marginal constraints $\mu_1, \dots, \mu_{n-1}$ are omitted; that is, only the first and last marginals μ_0, μ_n are prescribed. (One could similarly adapt the results to a case where some, but not all of the intermediate marginals are given.)

The primal space will be denoted by $\mathcal{M}^n(\mu_0, \mu_n)$ and consists of all martingale measures P on $\mathbb{R}^{n+1}$ such that $\mu_0 = P \circ (X_0)^{-1}$ and $\mu_n = P \circ (X_n)^{-1}$. To make the connection with the previous sections, we note that

$$\mathcal{M}^n(\mu_0, \mu_n) = \bigcup \mathcal{M}(\boldsymbol{\mu})$$

where the union is taken over all vectors $\boldsymbol{\mu} = (\mu_0, \mu_1, \dots, \mu_{n-1}, \mu_n)$ in convex order.

1.9.1 Polar Structure

We first characterize the polar sets of $\mathcal{M}^n(\mu_0, \mu_n)$. To that end, we introduce an analogue of the irreducible components.

Definition 1.9.1. Let $\mu_0 \leq_c \mu_n$ and let $(I_k, J_k) \subseteq \mathbb{R}^2$ be the corresponding irreducible domains in the sense of Proposition 1.2.3. The *n-step components* of $\mathcal{M}^n(\mu_0, \mu_n)$ are the sets[13]

(i) $I_k^n \times J_k$, where $k \geq 1$,

(ii) $I_0^{n+1} \cap \Delta_n$,

(iii) $I_k^t \times \{p\}^{n-t+1}$, where $p \in J_k \setminus I_k$ and $1 \leq t \leq n$, $k \geq 1$.

The characterization then takes the following form.

Theorem 1.9.2 (Polar Structure)**.** *Let $\mu_0 \leq_c \mu_n$. A Borel set $B \subseteq \mathbb{R}^{n+1}$ is $\mathcal{M}^n(\mu_0, \mu_n)$-polar if and only if there exist a μ_0-nullset N_0 and a μ_n-nullset N_n such that*

$$B \;\subseteq\; (N_0 \times \mathbb{R}^n) \;\cup\; (\mathbb{R}^n \times N_n) \;\cup\; \Big(\bigcup V_j\Big)^c$$

[13] A superscript m indicates the m-fold Cartesian product; Δ_n is the diagonal in $\mathbb{R}^{n+1}$.

where the union runs over all n-step components V_j of $\mathcal{M}^n(\mu_0, \mu_n)$.

It turns out that our previous results can be put to work to prove the theorem, by means of the following lemma which may be of independent interest.

Lemma 1.9.3. *Let $\mu \leq_c \nu$ be irreducible with domain (I, J) and let ρ be a probability concentrated on J. Then, there exists a probability $\mu \leq_c \theta \leq_c \nu$ satisfying $\theta \gg \rho$ such that $\mu \leq_c \theta$ and $\theta|_I \leq_c (\nu - \theta|_{J\setminus I})$ are both irreducible.*

Proof. Step 1. We first assume that $\rho = \delta_x$ for some $x \in J$ and show that there exists θ satisfying

$$\mu \leq_c \theta \leq_c \nu \quad \text{and} \quad \theta \gg \delta_x.$$

If ν has an atom at x, we can choose $\theta = \nu$. Thus, we may assume that $\nu(\{x\}) = 0$ and in particular that $x \in I$. Let a be the common barycenter of μ and ν and suppose that $x < a$. For all $b \in \mathbb{R}$ and $0 \leq c \leq \nu(\{b\})$, the measure

$$\nu_{b,c} := \nu|_{(-\infty,b)} + c\delta_b$$

satisfies $\nu_{b,c} \leq \nu$, and as $x < a$ there are unique b, c such that $\mathrm{bary}(\nu_{b,c}) = x$. Setting $\alpha = \nu_{b,c}$ and $\epsilon_0 = \alpha(\mathbb{R})$, we then have $\epsilon_0\delta_x \leq_c \alpha \leq \nu$, and a similar construction yields this result for $x \geq a$. The existence of such α implies that

$$\epsilon\delta_x \leq_{pc} \nu, \quad 0 \leq \epsilon \leq \epsilon_0$$

and thus the shadow $\mathcal{S}^\nu(\epsilon\delta_x)$ is well-defined. This measure is given by the restriction of ν to an interval (possibly including fractions of atoms at the endpoints); cf. [16, Example 4.7]. Moreover, the interval is bounded after possibly reducing the mass ϵ_0. Thus, for all $\epsilon < \epsilon_0$, the difference of potential functions

$$u_{\mathcal{S}^\nu(\epsilon\delta_x)} - u_{\epsilon\delta_x} \geq 0$$

vanishes outside a compact interval, and it converges uniformly to zero as $\epsilon \to 0$.

On the other hand, as $\mu \leq_c \nu$ is irreducible, the difference $u_\nu - u_\mu \geq 0$ is uniformly bounded away from zero on compact subsets of I and has nonzero derivative on $J \setminus I$. Together, it follows that

$$u_\nu - u_{\mathcal{S}^\nu(\epsilon\delta_x)} + u_{\epsilon\delta_x} \geq u_\mu \tag{1.9.1}$$

for small enough $\epsilon > 0$, so that

$$\theta := \nu - \mathcal{S}^\nu(\epsilon\delta_x) + \epsilon\delta_x$$

satisfies $\mu \leq_c \theta \leq_c \nu$; moreover, $\theta \gg \delta_x$ as $\nu(\{x\}) = 0$.

Step 2. We turn to the case of a general probability measure ρ on J. By Step 1, we can find a measure θ_x for each $x \in J$ such that

$$\mu \leq_c \theta_x \leq_c \nu \quad \text{and} \quad \theta_x \gg \delta_x.$$

The map $x \mapsto \theta_x$ can easily be chosen to be measurable (by choosing the ϵ for (1.9.1) in a measurable way). We can then define the probability measure

$$\theta'(A) := \int_J \theta_x(A)\rho(dx), \quad A \in \mathfrak{B}(\mathbb{R})$$

which satisfies $\mu \leq_c \theta' \leq_c \nu$. Moreover, we have $\theta' \gg \rho$; indeed, if $A \in \mathfrak{B}(\mathbb{R})$ is a θ'-nullset, then $\theta_x(A) = 0$ for ρ-a.e. x and thus $\rho(A) = 0$ as $\theta_x \gg \delta_x$.

Finally, $\theta := (\mu+\theta'+\nu)/3$ shares these properties. As $u_\mu < u_\nu$ on I due to irreducibility, we have $u_\mu < u_\theta < u_\nu$ on I and it follows that $\mu \leq_c \theta$ and $\theta|_I \leq_c (\nu - \theta|_{J\setminus I})$ are irreducible.

□

Lemma 1.9.4. *Let $\mu_0 \leq_c \mu_n$ and let π be a measure on $\mathbb{R}^{n+1}$ which is concentrated on an n-step component V of $\mathcal{M}^n(\mu_0, \mu_n)$ and whose first and last marginals satisfy*

$$\pi_0 \leq \mu_0, \quad \pi_n \leq \mu_n.$$

Then there exists $P \in \mathcal{M}^n(\mu_0, \mu_n)$ such that $P \gg \pi$.

Proof. If $V = I_0^{n+1} \cap \Delta_n$, then π must be an identical transport and we can take P to be any element of $\mathcal{M}(\mu_0, \mu_0, \dots, \mu_0, \mu_n)$. Thus, we may assume that V is of type (i) or (iii) in

Definition 1.9.1, and then, by fixing $k \geq 1$, that $\mu_0 \leq_c \mu_n$ is irreducible with domain (I, J).

Using Lemma 1.9.3, we can find intermediate marginals μ_t with

$$\mu_0 \leq_c \mu_1 \leq_c \cdots \leq_c \mu_{n-1} \leq_c \mu_n$$

such that $\mu_t \gg \pi_t$ for all $1 \leq t \leq n-1$, and each of the steps $\mu_{t-1} \leq_c \mu_t$, $1 \leq t \leq n$ has a single irreducible domain given by (I, J) as well as (possibly) a diagonal component on $J \setminus I$. We note that V is an irreducible component of $\mathcal{M}(\mu_0, \mu_1, \ldots, \mu_n)$ as introduced after Theorem 1.3.1.

Let $f_t = d\pi_t/d\mu_t$ be the Radon–Nikodym derivative of the marginal at date t. For $m \geq 1$, we define the measure $\pi^m \ll \pi$ by

$$\pi^m(dx_0, \ldots, dx_n) = 2^{-m} \left(\prod_{t=1}^{n-1} \mathbf{1}_{f_t(x_t) \leq 2^m} \right) \pi(dx_0, \ldots, dx_n).$$

Then, the marginals π_t^m satisfy the stronger condition $\pi_t^m \leq \mu_t$ for $0 \leq t \leq n$. Thus, we can apply Lemma 1.3.3 to $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_n)$ and the irreducible component V, to find $P^m \in \mathcal{M}(\boldsymbol{\mu}) \subseteq \mathcal{M}^n(\mu_0, \mu_n)$ such that $P^m \gg \pi^m$. Noting that $\sum_{m \geq 1} 2^{-m} \pi^m \gg \pi$, we see that $P := \sum_{m \geq 1} 2^{-m} P^m \gg \pi$ satisfies the requirements of the lemma. $\square$

Proof of Theorem 1.9.2. The result is deduced from Lemma 1.9.4 by following the argument in the proof of Theorem 1.3.1. $\square$

1.9.2 Duality

In this section we formulate a duality theorem for the transport problem with free intermediate marginals.

Definition 1.9.5. Let $f : \mathbb{R}^{n+1} \to [0, \infty]$. The *primal problem* is

$$\mathbf{S}^n_{\mu_0, \mu_n}(f) := \sup_{P \in \mathcal{M}^n(\mu_0, \mu_n)} P(f) \in [0, \infty]$$

and the dual problem is

$$\mathbf{I}^n_{\mu_0,\mu_n}(f) := \inf_{(\phi,\psi,H)\in\mathcal{D}^n_{\mu_0,\mu_n}(f)} \mu_0(\phi) + \mu_n(\psi) \in [0,\infty],$$

where $\mathcal{D}^n_{\mu_0,\mu_n}(f)$ consists of all triplets (ϕ,ψ,H) such that $(\phi,\psi) \in L^c(\mu_0,\mu_n)$ and $H = (H_1,\dots,H_n)$ is $\mathbb{F}$-predictable with

$$\phi(X_0) + \psi(X_n) + (H\cdot X)_n \geq f \quad \mathcal{M}^n(\mu_0,\mu_n)\text{-q.s.}$$

i.e. the inequality holds P-a.s. for all $P \in \mathcal{M}^n(\mu_0,\mu_n)$.

The analogue of Theorem 1.5.2 reads as follows.

Theorem 1.9.6 (Duality)**.** *Let $f : \mathbb{R}^{n+1} \to [0,\infty]$.*

(i) If f is upper semianalytic, then $\mathbf{S}^n_{\mu_0,\mu_n}(f) = \mathbf{I}^n_{\mu_0,\mu_n}(f) \in [0,\infty]$.

(ii) If $\mathbf{I}^n_{\mu_0,\mu_n}(f) < \infty$, there exists a dual optimizer $(\phi,\psi,H) \in \mathcal{D}^n_{\mu_0,\mu_n}(f)$.

The main step for the proof is again a closedness result. We shall only discuss the case where $\mu_0 \leq_c \mu_n$ is irreducible; the extension to the general case can be obtained along the lines of Section 1.4.

Proposition 1.9.7. *Let $\mu_0 \leq_c \mu_n$ be irreducible and let $f^m : \mathbb{R}^{n+1} \to [0,\infty]$ be a sequence of functions such that $f^m \to f$ pointwise. Moreover, let $(\phi^m,\psi^m,H^m) \in \mathcal{D}^n_{\mu_0,\mu_n}(f^m)$ be such that $\sup_m \mu_0(\phi^m) + \mu_n(\psi^m) < \infty$. Then there exist $(\phi,\psi,H) \in \mathcal{D}^n_{\mu_0,\mu_n}(f)$ such that*

$$\mu_0(\phi) + \mu_n(\psi) \leq \liminf_{m\to\infty} \mu_0(\phi^m) + \mu_n(\psi^m).$$

Proof. Let μ_t, $1 \leq t \leq n-1$ be such that $\boldsymbol{\mu} = (\mu_0,\dots,\mu_n)$ is in convex order and $\mu_{t-1} \leq_c \mu_t$ is irreducible for all $1 \leq t \leq n$; such μ_t are easily constructed by prescribing their potential functions. Setting $\boldsymbol{\phi}^m = (\phi^m, 0, \dots, 0, \psi^m)$ we have $(\boldsymbol{\phi}^m, H^m) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f^m)$ and can thus apply Proposition 1.4.21 to obtain $(\boldsymbol{\phi}, H) \in \mathcal{D}^g_{\boldsymbol{\mu}}(f)$. The construction in the proof of that proposition yields $\phi_t \equiv 0$ for $1 \leq t \leq n-1$. Therefore, $(\phi_0,\phi_n,H) \in \mathcal{D}^n_{\mu_0,\mu_n}(f)$ and

$$\mu_0(\phi_0) + \mu_n(\phi_n) = \boldsymbol{\mu}(\boldsymbol{\phi}) \leq \liminf_{m\to\infty} \boldsymbol{\mu}(\boldsymbol{\phi}^m) = \liminf_{m\to\infty} \mu_0(\phi^m) + \mu_n(\psi^m).$$

□

Proof of Theorem 1.9.6. On the strength of Proposition 1.9.7, the proof is analogous to the one of Theorem 1.5.2. □

1.9.3 Monotone Transport

The analogue of our result on left-monotone transports is somewhat degenerate: with unconstrained intermediate marginals, the corresponding coupling is the identical transport in the first $n-1$ steps and the (one-step) left-monotone transport in the last step. The full result runs as follows.

Theorem 1.9.8. *Let $P \in \mathcal{M}^n(\mu_0, \mu_n)$. The following are equivalent:*

(i) P is a simultaneous optimizer for $\mathbf{S}^n_{\mu_0,\mu_n}(f(X_0, X_t))$ for all smooth second-order Spence–Mirrlees functions f and $1 \leq t \leq n$.

(ii) P is concentrated on a left-monotone set $\Gamma \subset \mathbb{R}^{n+1}$ such that

$$\Gamma^{n-1} = \{(x, \dots, x) : x \in \Gamma^0\}.$$

(iii) For $0 \leq t \leq n-1$, we have $P \circ (X_t)^{-1} = \mu_0$ and $P \circ (X_t, X_n)^{-1}$ is the (one-step) left-monotone transport in $\mathcal{M}(\mu_0, \mu_n)$.

There exists a unique $P \in \mathcal{M}^n(\mu_0, \mu_n)$ satisfying (i)–(iii).

Proof. A transport P as in (iii) exists and is unique, because the identical transport between equal marginals and the left-monotone transport in $\mathcal{M}(\mu_0, \mu_n)$ exist and are unique; cf. Proposition 1.6.3. The equivalence of (ii) and (iii) follows from the same proposition and the fact that the only martingale transport from μ_0 to μ_0 is the identity.

Let $P \in \mathcal{M}^n(\mu_0, \mu_n)$ satisfy (i). In particular, P is then an optimizer for $\mathbf{S}^n_{\mu_0,\mu_n}(f(X_0, X_n))$, which by Proposition 1.6.3 implies that $P_{0n} = P \circ (X_0, X_n)^{-1}$ is the (one-step) left-monotone transport in $\mathcal{M}(\mu_0, \mu_n)$. For $t = 1, \dots, n-1$, P is an optimizer for $\mathbf{S}^n_{\mu_0,\mu_n}(-\mathbf{1}_{\{X_0 \leq a\}}|X_t - b|)$, for all $a, b \in \mathbb{R}$. This implies that P_{0t} transports $\mu_0|_{(-\infty,a]}$ to the minimal element of

$\{\theta : \mu_0|_{(-\infty,a]} \leq_c \theta \leq_{pc} \mu_n\}$ in the sense of the convex order, which is $\theta = \mu_0|_{(-\infty,a]}$. Therefore, P_{0t} must be the identical transport for $t = 1, \ldots, n-1$ and all but the last marginal are equal to μ_0.

Conversely, let $P \in \mathcal{M}^n(\mu_0, \mu_n)$ have the properties from (iii). Then, P is optimal for $\mathbf{S}^n_{\mu_0,\mu_n}(-\mathbf{1}_{\{X_0 \leq a\}}(X_t - b)^+)$ for all $1 \leq t \leq n$ and this can be extended to the optimality (i) for smooth second-order Spence–Mirrlees functions as in the proof of Theorem 1.7.12. $\square$

Chapter 2

Convergence to the Mean Field Game Limit: A Case Study

2.1 Introduction

Mean field games were introduced by [84, 85, 86] and [68, 69] to overcome the notorious intractability of n-player games. Two key simplifications are made. First, agents interact symmetrically through the empirical distribution of their states. Second, by formally letting $n \to \infty$, one passes to a representative agent whose actions do not affect this distribution because each individual agent becomes negligible. Thus, the mean field game is seen as an approximation of the n-player game for large n. We refer to the lecture notes [31] and the monographs [19, 33, 34] and their extensive references for further background.

In this paper, we conduct a case study of an n-player game of optimal stopping where multiple equilibria may occur naturally. We formulate an associated mean field game and highlight that certain mean field equilibria are limits of n-player equilibria while others are not, and study how to distinguish them. Equilibria that are not limit points are questionable from the point of view of applications, at least if they are motivated as "n-player games with large n."

Several ways of connecting n-player and mean field games have been studied in the literature. In many cases it is easier to establish the reverse direction, namely that a given mean field equilibrium induces an *approximate* Nash equilibrium in the n-player game for large n. This goes back to [69] and is by now established in some generality, see in particular [82] for diffusion control, [35] for games of timing or [38] for finite state games (but see also [30] for a counterexample in a degenerate case with absorption). It then follows, conversely, that mean field equilibria are limits of approximate n-player equilibria. However, we emphasize that approximate and actual Nash equilibria may look quite different, and in particular one cannot expect in general that there is a true Nash equilibrium in the proximity of an approximate one.

The convergence of n-player Nash equilibria to the mean field limit is often more delicate. The deep result of [32] shows convergence for a class of (closed-loop) games where agents choose drifts of diffusions. In their setting, the mean field game has a unique equilibrium as a consequence of the so-called monotonicity condition [84] which postulates that it is disadvantageous for agents' states to be close to one another. In a related but different (open-loop) framework, and without imposing uniqueness, [49] obtains convergence under the assumption that the limiting measure flow is deterministic. More comprehensively, [82] shows that n-player equilibria converge to a weak notion of mean field equilibria which can include mixtures of deterministic equilibria, for a general class of diffusion-control games. A corresponding result for games of timing is established in [35]. Most recently, [83] provides results along the lines of [82] for the closed-loop case. Convergence has also been shown in a number of more specific problems, for instance stationary mean field games [84], linear-quadratic problems [7] or a game of Poissonian control [96], among others. However, to the best of our knowledge, the question which mean field equilibria are limit points of (true) n-player equilibria has not been emphasized as such in the literature. We can mention the parallel work [39] on a two-state game: the game has unique n-player equilibria and these converge to a mean field equilibrium as expected; however, a second, less plausible mean field solution can appear for certain parameter values and this solution is not a limit. Another interesting parallel work [43] studies several approaches of selecting an equilibrium in a linear-quadratic mean field game with multiple equilibria, including the convergence of n-player equilibria. Different approaches are shown to select different equilibria.

From the perspective of mean field games, being a limit point of n-player equilibria can be seen as a stability property of equilibria with respect to the number of players. We are not aware of a systematic study in this direction (but see [27] for a recent investigation of a different stability property that is potentially related). Since mean field equilibria are often motivated as "large n" equilibria, it seems desirable to understand the phenomenon in some generality and at least establish sufficient conditions. A general formulation and investigation of this stability seems wide open at this time, whence our focus on a case study in the present paper.

2.1.1 Synopsis

We start by introducing an n-player game of optimal stopping inspired by [21, 35, 93] and the literature on bank-runs following [46]. In addition to their i.i.d. signals, players observe how many other players have already stopped. A crucial feature is that whenever an agent leaves the game, staying in the game becomes less attractive for the remaining agents. For instance, this may reflect that the bank is more likely to default if other clients withdraw their savings. In particular, the game satisfies the opposite of Lasry and Lions' monotonicity condition, or *strategic complementarity* in Economics terminology [28]. Indeed, the model exhibits a "flocking" or "herding" behavior where groups of agents can collectively decide to stop or not. We will see that these choices can naturally give rise to multiple equilibria; more precisely, they parametrize the full range of n-player equilibria.

Next, we review the mean field version of the game which was introduced in [93] without discussing the n-player game. Enhancing slightly a result of [93], mean field equilibria are described by a simple equation: for any equilibrium, the proportion $\rho(t)$ of agents that have stopped by time t is a zero of a deterministic function g_t on $[0, 1]$ as is Figure 2.1. More generally, any equilibrium $t \mapsto \rho(t)$ is characterized as an increasing, right-continuous selection of such zeros. In Figure 2.1, we can distinguish several types of zeros: increasing-transversal (i), tangential (t) and decreasing-transversal (d). These types are related to how concentrated the distribution of the agents' signals is in a neighborhood of the zero, relative to the strength of interaction. Intuitively, tangential solutions are delicate in that they may disappear if Figure 2.1 is perturbed, whereas the transversal solutions are stable in this sense.

We then turn to our main question and study which mean field equilibria are limits of n-player equilibria. Roughly, the main result is that

(i) Increasing-transversal solutions are limits of n-player equilibria,

(ii) decreasing-transversal solutions *fail* to be limits,

(iii) tangential solutions can but need not be limits.

Specifically, we first consider the minimal and maximal equilibria, corresponding to the left- and right-most solutions in Figure 2.1. The n-player game also has such extremal equilibria and these yield natural candidates for sequences converging to their mean field counterparts. After introducing appropriate notions for dynamic equilibria, we show that this convergence indeed holds, under the condition that the solutions are increasing-transversal (on a sufficiently large set of times t). However, we also find that if the minimal (say) solution is tangential, the minimal n-player equilibria can converge to a mixture of mean field equilibria and then the minimal mean field equilibrium may fail to be a proper limit. (The minimal and maximal solutions can be increasing-transversal or tangential, but never decreasing-transversal.) This also yields a novel example of how randomization can emerge in mean field games.

Second, we study the convergence to a general mean field equilibrium, possibly somewhere in the middle of Figure 2.1. In that case, there are no obvious candidates for the n-player approximations and more abstract arguments need to be used. We show by a fixed point construction that all increasing-transversal solutions are limits of n-player equilibria. Quite surprisingly however, ("strongly") decreasing-transversal solutions fail to be limits despite appearing stable in Figure 2.1. In fact, these solutions merely occur as parts of mixtures that are limits, and the weight within these mixtures can be bounded by a monotone function of the slope in Figure 2.1. It turns out that some fairly detailed asymp-

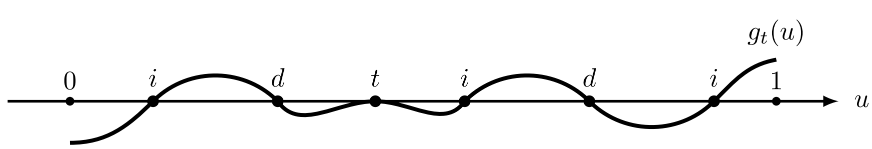


Figure 2.1: Types of mean field equilibria at a fixed time t

totic statistics, such as the expected number of n-player equilibria, can be analyzed in our model—which is unusual for mean field games.

The remainder of this paper is organized as follows. In Section 2.2, we introduce the game of optimal stopping. Section 2.3 describes the Nash equilibria of the n-player version and Section 2.4 covers the analogue for the mean field game. The results on the convergence to the minimal and maximal equilibria are relatively direct and established in Section 2.5, whereas the more abstract results on the convergence to general equilibria are reported in Section 2.6.

2.2 Description of the Game

Let $(I, \mathcal{I}, \lambda)$ be a probability space representing the agents; we shall be interested in the n-player case with a finite I and the mean field case with an atomless space. Let $(\Omega, \mathcal{G}, P)$ be another probability space, equipped with a right-continuous filtration $\mathbb{G} = (\mathcal{G}_t)_{t \in \mathbb{R}_+}$ and an exponentially distributed random variable $\mathcal{E}$ which is independent of $\mathbb{G}$.

Given an agent $i \in I$, let $\alpha^i \geq 0$ be a $\mathbb{G}$-progressively measurable process which is locally integrable and consider the random time

$$\theta^i = \inf \left\{ t : \int_0^t \alpha^i_s \, ds = \mathcal{E} \right\}.$$

As in [93], one may think of θ^i as the time when agent i expects the default of her bank. We fix a parameter $r \in \mathbb{R}$, interpreted as the interest rate paid by the bank (and assumed to be constant for simplicity). Following [93], we suppose that α^i is increasing[1] and that

$$\inf\{t : \alpha^i_t - r \geq 0\} < \infty \quad P\text{-a.s.} \tag{2.2.1}$$

Denoting by $\mathcal{T}$ the set of all $\mathbb{G}$-stopping times, we then consider the optimal stopping problem

$$\sup_{\tau \in \mathcal{T}} E\big[e^{r\tau} \mathbf{1}_{\{\theta^i > \tau\} \cup \{\theta^i = \infty\}}\big] \tag{2.2.2}$$

which we assume to have a finite value. Thus, if the default $\theta^i > \tau$, we may think of the

[1] Increase is to be understood in the non-strict sense throughout the paper.

agent as accruing the interest on an initial unit investment until τ, but losing everything if $\theta^i < \tau$. If the stopping time

$$\tau^i := \inf\{t : \alpha_t^i \geq r\} \in \mathcal{T} \tag{2.2.3}$$

is a.s. finite, then τ^i is optimal and in fact the minimal solution of (2.2.2); cf. [93, Lemma 2.1]. The solution is unique for instance if α^i is strictly increasing, but not in general. We assume that agents choose (2.2.3) in the case of non-uniqueness, which can be motivated e.g. as a preference for early stopping when other things are equal. This convention is not essential, but simplifies our exposition and allows us to focus on multiplicity of equilibria due to inherent game-theoretic aspects as it avoids ambiguity at the individual agents' level.

The processes α^i will depend on the proportion $\rho(t)$ of players who have already stopped, thus inducing an interaction among the agents. Since given ρ, the optimal stopping times are completely determined by (2.2.3), we shall simply say that an equilibrium is a process ρ which is $\mathbb{G}$-adapted and such that

$$\rho(t) = \lambda\{i : \tau^i \leq t\},$$

where it is tacitly assumed that the above set is λ-measurable.

2.3 The n-Player Game

In this section, we formulate the n-player version of the "toy model" mean field game in [93, Section 4]. Indeed, fix $n \in \mathbb{N}$ and take $I = \{1, \ldots, n\}$ to be a set with n elements, equipped with the normalized counting measure. Each player i observes an idiosyncratic signal $Y_t^i \geq 0$ which is right-continuous, progressively measurable, increasing and such that $\{Y^i\}_{i \in I}$ are pairwise i.i.d. with the common c.d.f.

$$y \mapsto F_t(y) := P\{Y_t^i \leq y\}.$$

Moreover, for a fixed interaction constant[2] $c > 0$,

$$\alpha^i(t) = Y^i_t + c\rho_n^{-i}(t), \quad \text{where} \quad \rho_n^{-i}(t) = \frac{\#\{j \neq i : \tau^j \leq t\}}{n}$$

is the fraction of other players[3] (from the perspective of i) that have already stopped, according to $(\tau^j)_{j\neq i}$. Specializing from the previous section, an n-player equilibrium boils down to the process $\rho_n(t) = \#\{j : \tau^j \leq t\}/n$ where τ^j are as in (2.2.3). In particular, if ρ_n is an equilibrium and (t, ω) is such that $\rho_n(t)(\omega) = k/n$, then as the stopping times satisfy $\tau^i = \inf\{t : \alpha^i(t) \geq r\}$, we must have[4]

$$\#\{Y^i_t(\omega) + c\,\frac{k-1}{n} \geq r\} = k \quad \text{and} \quad \#\{Y^i_t(\omega) + c\,\frac{k}{n} < r\} = n - k. \tag{2.3.1}$$

This condition is also sufficient, in the sense made precise in Remark 2.3.5.

Next, we sketch the structure of all equilibria $\rho_n(t) = \#\{i : \tau^i \leq t\}/n$ of this game by a recursive construction, starting with $K = \emptyset$.

1. Suppose that at a given stopping time t_0, a group $K \subsetneq I$ of agents has already stopped. Then every remaining agent $i \notin K$ examines her criterion

$$\theta^i_K = \inf\{t : \, Y^i_t + c\,\frac{\#K}{n} \geq r\}.$$

If $\theta^i_K \leq t_0$, then player i must stop immediately. We add i to the set K and repeat Step 1 until no further players are forced to stop. (By the monotonicity in $\#K$, it does not matter in which order the agents are processed.)

2. Beyond individual players forced to stop, a group $J \subseteq K^c$ of agents may be able to

[2] We could more generally consider processes α^i which are nonlinear functions of Y^i and ρ^{-i} and possibly a common noise, as in [93]. However, the increased generality does not seem to lead to additional insights regarding the main questions of this paper, so we have chosen to use the simplified "toy model" in our exposition. The constant c could in fact be normalized to 1 by changing Y^i and r, but we find it useful to represent the strength of interaction explicitly.

[3] Once again, we have decided to exclude player i in order to focus on the game-theoretic aspect of multiplicity. If player i considers her own action; i.e., uses ρ instead of ρ^{-i}, non-uniqueness can occur without other agents' involvement simply because of the direct feedback on the state process.

[4] We will often abbreviate $\#\{i \in I : \dots\}$ to $\#\{\dots\}$ in what follows.

"coordinate" and stop together.[5] Indeed, suppose that

$$\theta_K^J = \inf\{t : \, Y_t^i + c\,\frac{\#K + \#J - 1}{n} \geq r\}$$

satisfies $\theta_K^J \leq t_0$ for all $i \in J$. Then it is optimal for all these agents to stop as a group, and they may or may not "choose" to do so. If they stop, we add J to K and repeat the procedure starting with Step 1.

3. After all remaining groups of agents have decided whether to stop at time t_0, we increment time until there exists a group or individual agent wanting to stop, and start again at Step 1.

The multiplicity of equilibria of this game arises because of the choices taken by the groups J in Step 2, as well as the order in which the groups are processed. Next, we describe two of these equilibria in detail. The first one is the minimal equilibrium and corresponds to groups J in Step 2 always choosing not to stop. This is equivalent to all players remaining in the game until their own optimality criterion forces them to quit.

Proposition 2.3.1. *There exists an n-player equilibrium ρ_n^m such that*

$$\rho_n^m(t) = \frac{k}{n} \iff \begin{cases} \#\{Y_t^i + c\,\dfrac{k}{n} \geq r\} = k \\ \#\{Y_t^i + c\,\dfrac{k-l}{n} \geq r\} \geq k - l + 1, \quad 1 \leq l \leq k. \end{cases} \tag{2.3.2}$$

This equilibrium is minimal; i.e., $\rho_n^m(t) \leq \rho_n(t)$ for any n-player equilibrium ρ_n.

Proof. The construction is iterative. Given a set $K \subsetneq I$ corresponding to players who have already stopped, we can consider for all $i \notin K$ the stopping times

$$\theta_K^i = \inf\{t : \, Y_t^i + c\,\frac{\#K}{n} \geq r\}$$

with the corresponding order statistics $\theta_K^{(1)} \leq \theta_K^{(2)} \leq \dots$. We define $\theta_K = \theta_K^{(1)}$ and $i_K = (1)$. We note that agent i must stop at θ_K^i, even if no further agents $j \notin K$ choose to stop, and that i_K is the first of the agents $i \notin K$ subject to this event.

[5]While we are using suggestive language here, it should be noted that these are simply different configurations which may be equilibria. We are not trying to model a mechanism how players "find" an equilibrium.

To define the equilibrium, start with $K_0 = \emptyset$ and set $\tau^i = \theta_{K_0} \equiv \theta_{K_0}^{(1)}$ on $\{i = i_{K_0}\}$. Next, set $K_1 = \{i_{K_0}\}$ and $\tau^i = \max\{\theta_{K_1}, \theta_{K_0}\}$ on $\{i = i_{K_1}\}$, and continue inductively setting $K_k = K_{k-1} \cup \{i_{K_{k-1}}\}$ and $\tau^i = \max\{\theta_{K_k}, \tau^{i_{K_{k-1}}}\}$ on $\{i = i_{K_k}\}$ for $k = 2, \dots, n-1$. (The maximum needs to be taken since all the α^j are increased after player $i_{K_{k-1}}$ stops.)

Setting $\rho_n^m(t) = \#\{i : \tau^i \leq t\}/n$, we have by construction that ρ_n^m is an equilibrium with corresponding optimal stopping times (τ^i) and that (2.3.2) holds.

To see the minimality, let ρ_n be any n-player equilibrium and consider (t, ω) such that $\rho_n(t)(\omega) = k/n$. Let k' be such that $\rho_n^m(t)(\omega) = k'/n$. If we had $k' > k$, then (2.3.2) would imply $\#\{Y_t^i(\omega) + c\,\frac{k}{n} \geq r\} \geq k+1$ and hence $\#\{Y_t^i(\omega) + c\,\frac{k}{n} < r\} \leq n-k-1$, a contradiction to (2.3.1). Thus, $k' \leq k$ and we have shown that $\rho_n^m \leq \rho_n$. □

Remark 2.3.2. Let ρ be an n-player equilibrium and t_0 a stopping time. There exists an equilibrium which is minimal among all n-player equilibria ϱ such that $\varrho = \rho$ on $[0, t_0]$. Indeed, it is obtained by agents stopping as in ρ until t_0, whereas from t_0 onwards we apply the construction in the proof of Proposition 2.3.1 starting with $K = \{i : \tau^i \leq t_0\}$. We call this ϱ the *minimal extension* of ρ after t_0.

The second extremal equilibrium is maximal and corresponds to players coordinating their actions such as to stop as early as possible. As seen in the construction below, this is equivalent to all players constantly seeking (maximally large) groups of collaborators so that immediate simultaneous stopping is optimal for all agents in the group.

Proposition 2.3.3. *There exists an n-player equilibrium ρ_n^M such that*

$$\rho_n^M(t) = \frac{k}{n} \iff \begin{cases} \#\{Y_t^i + c\,\dfrac{k-1}{n} \geq r\} = k \\ \#\{Y_t^i + c\,\dfrac{k+l-1}{n} \geq r\} \leq k+l-1, \quad 1 \leq l \leq n-k. \end{cases} \tag{2.3.3}$$

This equilibrium is maximal; i.e., $\rho_n^M(t) \geq \rho_n(t)$ for any n-player equilibrium ρ_n.

Proof. Given a set $K \subsetneq I$ of size $k = \#K$ corresponding to players who have already stopped, we can consider for $1 \leq l \leq n-k$ the stopping times

$$\theta_K^l = \inf\{t : \#\{i \notin K : Y_t^i + c\,\frac{k+l-1}{n} \geq r\} \geq l\};$$

intuitively, this is the first time an additional group J of $\#J = l$ agents can collectively stop. If $\theta_K^{(1)} \leq \cdots \leq \theta_K^{(n-k)}$ are the corresponding order statistics (ties are split by assigning the lower rank to the larger index l), pick $l = (1)$ and let $J = J(K)$ be the set of $i \notin K$ such that $\{Y^i_{\theta^l_K}\}_{i \in J}$ are the l largest elements in $\{Y^i_{\theta^l_K}\}_{i \in K^c}$; we think of J as the l most pessimistic agents remaining at time θ^l_K and denote $\theta_K := \theta^l_K$.

To define the equilibrium, start with $K_0 = \emptyset$ and set $\tau^i = \theta_\emptyset$ for $i \in J(\emptyset)$. Next, set $K_1 = J(\emptyset)$ and $\tau^i = \theta_{K_1}$ for $i \in J(K_1)$, and continue inductively with $K_2 = J(K_1) \cup K_1$. Setting $\rho_n^M(t) = \#\{i : \tau^i \leq t\}/n$, we have by construction that ρ_n^M is an equilibrium with corresponding optimal stopping times (τ^i) and that (2.3.3) holds.

To see the maximality, let ρ_n be any n-player equilibrium and consider (t, ω) such that $\rho_n(t)(\omega) = k/n$. Again, ρ_n must satisfy (2.3.1). Let k' be such that $\rho_n^M(t)(\omega) = k'/n$. If we had $k' < k$, then (2.3.3) would imply that $\#\{Y^i_t(\omega) + c\,\frac{k-1}{n} \geq r\} \leq k-1$, contradicting (2.3.1). $\square$

The following observations will be used in Section 2.6 when we construct n-player equilibria converging to a given mean field equilibrium.

Remark 2.3.4. (i) Consider n-player equilibria ρ and ρ', a stopping time t_0 and assume that $\rho(t_0) \leq \rho'(t_0)$. Then there exists an n-player equilibrium ϱ such that

$$\varrho \mathbf{1}_{[0,t_0)} = \rho \mathbf{1}_{[0,t_0)} \quad \text{and} \quad \varrho \mathbf{1}_{[t_0,\infty)} = \rho' \mathbf{1}_{[t_0,\infty)}.$$

Indeed, let I_0 be the set of agents that have stopped by time t_0 in equilibrium ρ and let I_1 be the analogue for ρ'. By (2.2.3) we necessarily have $I_0 \subseteq I_1$. The equilibrium ϱ is obtained by following the stopping times of ρ on $[0, t_0)$. At t_0, all agents in the group $J = I_1 \setminus I_0$ stop (and this must be optimal as ρ' is an equilibrium). After that the remaining agents act as in ρ'.

(ii) Extending the above, consider n-player equilibria ρ and ρ', stopping times $t_0 \leq t_1$ and assume that $\rho(t_0) \leq \rho'(t_1)$. Then there exists an n-player equilibrium ϱ such that

$$\varrho \mathbf{1}_{[0,t_0)} = \rho \mathbf{1}_{[0,t_0)} \quad \text{and} \quad \varrho \mathbf{1}_{[t_1,\infty)} = \rho' \mathbf{1}_{[t_1,\infty)}. \tag{2.3.4}$$

Indeed, let ρ_1 be the minimal extension of ρ after t_0 (cf. Remark 2.3.2). Let I_0 be the set of agents that have stopped by time t_0 in equilibrium ρ and let I_1 be the set of agents that have stopped by time t_1 in equilibrium ρ'. Again, we observe that $I_0 \subseteq I_1$, due to (2.2.3) and the increase of Y^i. Moreover, I_1 must include all agents that stop in the construction of the minimal extension on $[t_0, t_1]$. As a result, $\rho_1(t_1) \leq \rho'(t_1)$, and now the claim follows by applying (i).

(iii) A last generalization is that when $\rho(t_0) \leq \rho'(t_1)$ merely holds on some set $A \in \mathcal{G}_{t_1}$, then we can still construct an n-player equilibrium ϱ satisfying (2.3.4) on A. Indeed, ϱ is found as in (ii) except that on A^c, agents continue to stop according to ρ_1 after t_1.

Remark 2.3.5. (i) The necessary condition (2.3.1) is sufficient in the following sense. Fix n and a stopping time t_0, and suppose there exists an $\mathcal{G}_{t_0}$-measurable random variable k satisfying (2.3.1) at t_0; i.e.,

$$\#\{Y^i_{t_0} + c\,\frac{k-1}{n} \geq r\} = k \quad \text{and} \quad \#\{Y^i_{t_0} + c\,\frac{k}{n} < r\} = n - k.$$

Then there exists an n-player equilibrium ϱ such that $\varrho(t_0) = k/n$.

To construct ϱ, let agents stop as in the minimal equilibrium ρ^m_n up to time t_0. By the argument at the end of the proof of Proposition 2.3.1, we must have $\rho^m_n(t_0) \leq k/n$. At t_0, all remaining agents i with $Y^i_{t_0} + c\,\frac{k-1}{n} \geq r$ stop, so that $\rho(t_0) = k/n$. After that, the remaining agents follow the construction in the proof of Proposition 2.3.1 starting with $K = \{i : \tau^i \leq t_0\}$.

(ii) A variant of this holds when (2.3.1) is satisfied on some set $A \in \mathcal{G}_{t_0}$, with the conclusion that $\varrho(t_0) = k/n$ holds only on A. Indeed, we construct ϱ as above on A, whereas on A^c we use ρ^m_n.

(iii) For later use, we observe that if this construction is applied for two times $t_0 \leq t_1$ and corresponding random variables $k_0 \leq k_1$, the resulting equilibria satisfy $\varrho_0 \leq \varrho_1$.

2.4 The Mean Field Game

The game considered in this section is the "toy model" mean field game of [93, Section 4]. Indeed, $(I, \mathcal{I}, \lambda)$ is an atomless probability space and we work on a so-called Fubini extension

$(I \times \Omega, \Sigma, \mu)$ of the product $(I \times \Omega, \mathcal{I} \times \mathcal{G}, \lambda \times P)$; see [93, Section 3]. For each $i \in I$, let $Y_t^i \geq 0$ be a right-continuous, increasing, $\mathbb{G}$-progressively measurable process such that for each $t \geq 0$, $(i, \omega) \mapsto Y_t^i(\omega)$ is Σ-measurable and $Y_t^i,\ i \in I$ are λ-essentially pairwise i.i.d.; see also [93, Definition 3.1]. Working on a Fubini extension ensures that such processes exist, as well as the validity of an Exact Law of Large Numbers. In all that follows, we assume that the c.d.f. $y \mapsto F_t(y) = P\{Y_t^i \leq y\}$ is continuous.

Since λ is atomless, each individual agent has zero mass and hence does not influence the state process $\rho(t) = \lambda\{i : \tau^i \leq t\}$. In particular, we do not distinguish ρ and ρ^{-i} and simply set $\alpha^i(t) = Y_t^i + c\rho(t)$. We recall that ρ is an equilibrium if $\rho(t) = \lambda\{i : \tau^i \leq t\}$ where τ^i is as in (2.2.3) for λ-a.e. $i \in I$. Such a process may be random (see also [93]). However, as common in the mean field game literature, we pay special attention to equilibria which are deterministic due to the infinite number of players.[6] The following is an improved version of [93, Proposition 4.1] with necessary and sufficient conditions.

Proposition 2.4.1. *A real function $\rho : \mathbb{R}_+ \to [0, 1]$ is a mean field game equilibrium if and only if it is increasing, right-continuous and*

$$\rho(t) + F_t(r - c\rho(t)) = 1, \quad t \geq 0. \tag{2.4.1}$$

Proof. Suppose that ρ is a mean field game equilibrium, then ρ is clearly increasing. Since $Y_t^i,\ i \in I$ are λ-essentially pairwise i.i.d., the Exact Law of Large Numbers (e.g., [93, Section 3]) states that $\lambda\{i : Y_t^i \leq u\} = F_t(u)$ for all u. Using also (2.2.3) and that $y \mapsto F_t(y)$ is continuous, we have

$$\rho(t) = \lambda\{i : \tau^i \leq t\} = \lambda\{i : Y_t^i + c\rho(t+) \geq r\} = 1 - F_t(r - c\rho(t+)). \tag{2.4.2}$$

Recall that Y^i has right-continuous paths. Using again the continuity of F_t, this implies that

$$(t, u) \mapsto F_t(r - cu) \text{ is jointly right-continuous.} \tag{2.4.3}$$

It follows that $t \mapsto 1 - F_t(r - c\rho(t+))$ is right-continuous, and thus the left-hand side of (2.4.2)

[6] Note that the key message of this paper, namely that some mean field equilibria are not limits of n-player equilibria, is only amplified if more mean field equilibria are considered.

must also be right-continuous. That is, $\rho(t) = \rho(t+)$, and then (2.4.2) becomes (2.4.1).

Conversely, suppose that ρ is a function with the stated properties. Defining the corresponding optimal stopping times τ^i as in (2.2.3), the Exact Law of Large Number shows that

$$\lambda\{i : \tau^i \leq t\} = \lambda\{i : Y_t^i + c\rho(t) \geq r\} = 1 - F_t(r - c\rho(t)) = \rho(t);$$

that is, ρ is an equilibrium. $\square$

The following notions will be crucial in determining the convergence to the mean field limit.

Definition 2.4.2. Fix $t \geq 0$. A solution $u \in [0,1]$ of $u + F_t(r - cu) = 1$ is called *left-increasing-transversal* (or left-transversal for short) if

$$\text{for all } \varepsilon > 0 \text{ there is } u' \in (u - \varepsilon, u) \text{ such that } u' + F_t(r - cu') < 1 \tag{2.4.4}$$

and *right-increasing-transversal* (or right-transversal) if

$$\text{for all } \varepsilon > 0 \text{ there is } u' \in (u, u + \varepsilon) \text{ such that } u' + F_t(r - cu') > 1. \tag{2.4.5}$$

It is called *increasing-transversal* if both (2.4.4) and (2.4.5) hold, and *decreasing-transversal* if these hold with the inequality signs reversed.

For instance, in Figure 2.2, u^m is left-increasing-transversal and u^{mrt}, u^M are right-increasing-transversal, but only u^{Mlt} is increasing-transversal. A decreasing-transversal solution is also depicted. Next, we introduce a quartet of solutions that will be important in Section 2.5.

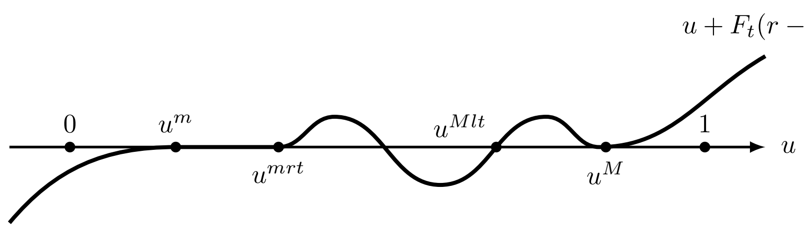


Figure 2.2: Solutions u^m, u^{mrt}, u^{Mlt} and u^M

Lemma 2.4.3. *Fix $t \geq 0$. The equation $u + F_t(r - cu) = 1$ has a minimal solution $u^m \in [0,1]$, a maximal solution $u^M \in [0,1]$, a minimal right-transversal solution $u^{mrt} \in [0,1]$, and a maximal left-transversal solution $u^{Mlt} \in [0,1]$.*

Proof. Since $G(u) := u + F_t(r - cu) < 1$ for $u < 0$ and $G(u) > 1$ for $u > 1$, the existence of u^m and u^M is immediate from the continuity of G. The fact that $G(u) < 1$ for all $u < u^m$ entails that u^m is left-transversal, and since it follows directly from the definition that the set of left-transversal solutions is stable under increasing limits, it follows that u^{Mlt} exists. The argument for u^{mrt} is similar. □

As illustrated in Figure 2.2, these four solutions may be distinct, and while u^m is automatically left-transversal, it can happen that u^{mrt} is not. Similarly for u^M and u^{Mlt}. We can also note that $u^{mrt} \leq u^{Mlt}$ may fail, say if the graph is replaced by a flat stretch on $[u^m, u^M]$. But in more generic cases, and in particular whenever u^m and u^M are not local extrema, the quartet describes at most two distinct solutions $u^m = u^{mrt} \leq u^{Mlt} = u^M$ and these are then increasing-transversal.

In view of Lemma 2.4.3 we may define, given $t \geq 0$,

$$\rho^m(t) = u^m, \quad \rho^M(t) = u^M, \quad \rho^{mrt}(t) = u^{mrt}, \quad \rho^{Mlt}(t) = u^{Mlt}. \tag{2.4.6}$$

Using the increase of Y_t and (2.4.3), one can check that $\rho^m, \rho^M, \rho^{mrt}, \rho^{Mlt}$ are increasing, ρ^M and ρ^{mrt} are right-continuous, and ρ^m and ρ^{Mlt} are left-continuous (but not continuous in general).

Corollary 2.4.4. *(i) If $\rho : \mathbb{R}_+ \to [0,1]$ is any increasing function such that (2.4.1) holds, then $\rho(t+)$ is an equilibrium.*

(ii) The functions $t \mapsto \rho^m(t+)$ and $t \mapsto \rho^M(t)$ are the minimal and maximal equilibria of the mean field game; i.e., they are equilibria and any other equilibrium ρ satisfies $\rho^m(t+) \leq \rho(t) \leq \rho^M(t)$ for all $t \geq 0$.

Proof. (i) If ρ is any increasing function such that (2.4.1) holds, then the joint right-continuity in (2.4.3) implies that $\rho(t+) + F_t(r - c\rho(t+)) = 1$ for all $t \geq 0$. It now follows from Proposition 2.4.1 that $\rho(t+)$ is an equilibrium.

(ii) Both $\rho^m(t+)$ and $\rho^M(t)$ are equilibria by (i). If ρ is any equilibrium, then it is necessarily right-continuous by Proposition 2.4.1 and thus $\rho^m \leq \rho \leq \rho^M$ implies $\rho^m(t+) \leq \rho(t) \leq \rho^M(t)$ for all $t \geq 0$. □

2.5 Convergence to Extremal Equilibria

The main goal of the last two sections is to understand which mean field equilibria are limits of n-player equilibria. In brief, we will see that mean field equilibria described by increasing-transversal solutions of (2.4.1) (on a sufficiently large sets of times t) are such limits, whereas other equilibria need not be proper limits of n-player equilibria; they merely occur as parts of mixtures which are limits.

In this section, we focus on the convergence to the minimal and maximal mean field equilibria; the less straightforward interior case is treated in the next section. As a first step, we relate limits of arbitrary n-player equilibria to mean field equilibria at a fixed time. We will see in Example 2.5.8 that such limits need not be deterministic mean field equilibria as defined in the preceding section, hence the following result relates limits to mixtures of equilibria. This is in line with the results of [35, 82] stating that n-player equilibria converge to "weak" equilibria of the mean field game, while also illustrating that randomization can indeed occur in a quite natural example.

Given a closed set $A \subseteq \mathbb{R}$, we say that a sequence (ξ_n) of random variables is *asymptotically concentrated* on A if $\lim_{n\to\infty} P(\xi_n \in A_\varepsilon) = 1$ for all $\varepsilon > 0$, where $A_\varepsilon = \{x \in \mathbb{R} : d(x, A) < \varepsilon\}$ is the open ε-neighborhood of A. When (ξ_n) is uniformly bounded, as it will be the case below, this is equivalent to any weak cluster point of (ξ_n) being concentrated on A. Moreover, for $t \geq 0$, we denote the solutions of (2.4.1) by

$$\mathcal{U}(t) = \{u \in [0, 1] : u + F_t(r - cu) = 1\}.$$

Proposition 2.5.1. *Fix $t \geq 0$ and let $(\rho_n)_{n\geq 1}$ be a sequence of n-player equilibria. Then $\rho_n(t)$ is asymptotically concentrated on $\mathcal{U}(t)$.*

Proof. We first show that for any interval $[u_0, u_1] \subseteq [0, 1]$ such that $u \mapsto u + F_t(r - cu)$ is strictly smaller than 1 on $[u_0, u_1]$,

$$P(u_0 + \varepsilon' \leq \rho_n(t) \leq u_1 - \varepsilon') \to 0 \quad \text{for all} \quad \varepsilon' > 0. \tag{2.5.1}$$

Indeed, let $u_0 < u_1$ be as above. By increasing the value of u_1 if necessary, we may assume without loss of generality that $u \mapsto u + F_t(r - cu)$ attains its maximum over $[u_0, u_1]$ at u_1. Given $0 < \varepsilon < u_1 - u_0$, we can then choose by continuity some $u \in (u_1 - \varepsilon, u_1)$ such that

$$u' + F_t(r - cu') \leq u + F_t(r - cu) < 1 \quad \text{for all} \quad u_0 \leq u' \leq u. \tag{2.5.2}$$

Furthermore, setting

$$\varepsilon_n(x) = \frac{\#\{Y_t^i + cx \geq r\}}{n} - (1 - F_t(r - cx)), \quad x \in \mathbb{R}$$

and $\varepsilon_n = \sup_{x\in\mathbb{R}}\{|\varepsilon_n(x)|\}$, we have $\varepsilon_n \to 0$ a.s. by the uniform convergence in the Glivenko–Cantelli theorem. Let $X_i = \mathbf{1}_{\{Y_t^i + cu \geq r\}}$, then

$$\frac{X_1 + \cdots + X_n}{n} = 1 - F_t(r - cu) + \varepsilon_n(u). \tag{2.5.3}$$

Denote by $[x]$ the largest integer $k \leq x$. For any $[u_0 n] + 1 \leq l \leq [un]$, let $Z_i^l = \mathbf{1}_{\{Y_t^i + c\,\frac{l}{n} \geq r\}}$, then similarly

$$\frac{Z_1^l + \cdots + Z_n^l}{n} = 1 - F_t(r - c\,\tfrac{l}{n}) + \varepsilon_n(\tfrac{l}{n}).$$

On the event $\{Z_1^l + \cdots + Z_n^l = l\}$, we then have

$$1 + \varepsilon_n(\tfrac{l}{n}) = \frac{l}{n} + F_t(r - c\,\tfrac{l}{n}) \leq u + F_t(r - cu)$$

by (2.5.2) and thus

$$\frac{X_1 + \cdots + X_n}{n} = 1 - F_t(r - cu) + \varepsilon_n(u) \leq u - \varepsilon_n(\tfrac{l}{n}) + \varepsilon_n(u) \leq u + 2\varepsilon_n.$$

Combining this observation with (2.3.1), we have for all $[u_0 n]+1 \le l \le [un]$ that

$$\begin{aligned}\Big\{\rho_n(t) = \frac{l}{n}\Big\} &\subseteq \Big\{\#\{Y_t^i + c\,\tfrac{l}{n} \ge r\} = l\Big\} \\ &= \Big\{\frac{Z_1^l + \cdots + Z_n^l}{n} = \frac{l}{n}\Big\} \\ &\subseteq \Big\{\frac{X_1 + \cdots + X_n}{n} \le u + 2\varepsilon_n\Big\}.\end{aligned}$$

Hence,

$$\Big\{\frac{[u_0 n]+1}{n} \le \rho_n(t) \le \frac{[un]}{n}\Big\} \subseteq \Big\{\frac{X_1 + \cdots + X_n}{n} \le u + 2\varepsilon_n\Big\}$$

and thus

$$\begin{aligned}P\Big(\frac{[u_0 n]+1}{n} \le \rho_n(t) \le \frac{[un]}{n}\Big) &\le P\Big(\frac{X_1 + \cdots + X_n}{n} \le u + 2\varepsilon_n\Big) \\ &= P\Big(u + F_t(r-cu) \ge 1 - 2\varepsilon_n + \varepsilon_n(u)\Big) \to 0\end{aligned}$$

by (2.5.3) and (2.5.2). Since $\varepsilon > 0$ was arbitrary, this shows (2.5.1).

In a symmetric way, one can show the analogue of (2.5.1) for intervals where $u \mapsto u + F_t(r-cu)$ is strictly larger than 1. Since for any $\varepsilon > 0$ the complement of $\mathcal{U}(t)_\varepsilon$ consists of finitely many intervals of one of these two types, the claim follows. □

Next, we narrow down the asymptotic support for the minimal and maximal n-player equilibria ρ_n^m and ρ_n^M. We will see in Section 2.5.1 that the following result is optimal and the limiting support is not a singleton in general. We recall the notation introduced in (2.4.6).

Lemma 2.5.2. *Fix $t \ge 0$.*

(i) The minimal n-player equilibrium $\rho_n^m(t)$ is asymptotically concentrated on $[\rho^m(t), \rho^{mrt}(t)] \cap \mathcal{U}(t)$.

(ii) The maximal n-player equilibrium $\rho_n^M(t)$ is asymptotically concentrated on $[\rho^{Mlt}(t), \rho^M(t)] \cap \mathcal{U}(t)$.

Proof. (i) In view of Proposition 2.5.1 and the definition of $\rho^m(t)$, it suffices to show that

$$P(\rho_n^m(t) \ge \rho^{mrt}(t) + \varepsilon') \to 0 \quad \text{for all} \quad \varepsilon' > 0. \tag{2.5.4}$$

Let $\varepsilon > 0$. As $\rho^{mrt}(t)$ is right-transversal we can find $u \in (\rho^{mrt}(t), \rho^{mrt}(t) + \varepsilon)$ such that $1 - F_t(r - cu) < u$. For n large enough, we then have $\rho^{mrt}(t) < [un]/n \leq u$. Let $X_i = \mathbf{1}_{\{Y_t^i + cu \geq r\}}$, then

$$\frac{X_1 + \cdots + X_n}{n} \to EX_i = 1 - F_t(r - cu) \quad \text{a.s.}$$

by the Law of Large Numbers. Hence,

$$\frac{X_1 + \cdots + X_n}{n} - \frac{[un]}{n} \to 1 - F_t(r - cu) - u < 0 \quad \text{a.s.}$$

Using also (2.3.2), we conclude that

$$\begin{aligned} P(\rho_n^m(t) \geq u) &\leq P\Big(\rho_n^m(t) \geq \frac{[un]}{n}\Big) \\ &\leq P\Big(\#\big\{Y_t^i + c\,\frac{[un]}{n} \geq r\big\} \geq [un]\Big) \\ &\leq P\Big(\frac{\#\{Y_t^i + cu \geq r\}}{n} \geq \frac{[un]}{n}\Big) \\ &= P\Big(\frac{X_1 + \cdots + X_n}{n} - \frac{[un]}{n} \geq 0\Big) \to 0. \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, the above implies (2.5.4).

(ii) The arguments are similar to (i) and therefore omitted. □

Next, we introduce an appropriate notion of convergence for dynamic equilibria as required for our main results. Note that given an increasing function, its right- and left-continuous limits (and all functions between these) differ only by the allocation of the function value at the (countably many) jumps. The fact that mean field equilibria are right-continuous, cf. Proposition 2.4.1, reflects the fact that agents stopping at time t are counted as having left the game at time t, whereas left-continuity would correspond to counting them as leaving immediately after t. Since this difference is not fundamental, it seems reasonable to consider limits "up to taking right-continuous versions." This has been accomplished by notions of so-called Fatou convergence, e.g. [80, 113], in other areas of stochastic analysis.

For increasing functions φ_n, φ on $\mathbb{R}_+$, we have that
$(\liminf_n \varphi_n)(t+) = (\limsup_n \varphi_n)(t+) = \varphi(t+)$ holds for all $t \in \mathbb{R}_+$ if and only if $\lim \varphi_n(t) =$

$\varphi(t)$ for all t in a dense subset $D \subseteq \mathbb{R}_+$. This motivates the following.

Definition 2.5.3. A sequence $(\rho_n)_{n\geq 1}$ of n-player equilibria *Fatou converges in probability* to a mean field equilibrium ρ if there exists a dense set $D \subseteq \mathbb{R}_+$ such that $\rho_n(t) \to \rho(t)$ in probability for all $t \in D$.

We note that by a diagonalization procedure, Fatou convergence in probability implies Fatou convergence a.s. along a subsequence (n_k), where the a.s. convergence is defined by direct analogy to the above. In particular, it then follows that the right-continuous versions of $\liminf_k \rho_{n_k}$ and $\limsup_k \rho_{n_k}$ coincide with ρ a.s.

With these notions in place, we can establish the convergence of extremal equilibria in the increasing-transversal case. (Note that the extremal equilibria cannot be decreasing-transversal; they are either increasing-transversal or tangential.)

Theorem 2.5.4. *Suppose that for all t in a dense subset $D \subseteq \mathbb{R}_+$, the minimal solution $u \in [0,1]$ of $u+F_t(r-cu) = 1$ is increasing-transversal. Then the minimal n-player equilibria ρ_n^m Fatou converge in probability to the minimal mean field equilibrium as $n \to \infty$.*

The analogous assertion holds for the maximal equilibria ρ_n^M.

Proof. By the hypothesis, $\rho^m(t) = \rho^{mrt}(t)$ for $t \in D$. Thus, Lemma 2.5.2 implies that $\lim \rho_n^m(t) = \rho^m(t) = \rho^{mrt}(t)$ in probability for $t \in D$. The analogue holds for ρ_n^M. □

Next, we discuss the transversality condition in more detail. In fact, if uniqueness holds for the mean field game, the condition is automatically satisfied and we conclude the following.

Corollary 2.5.5. *The following are equivalent:*

(i) the mean field game has a unique equilibrium ρ,

(ii) the equation $u + F_t(r - cu) = 1$, $u \in [0,1]$ has a unique solution for a dense set of $t \in \mathbb{R}_+$.

In that case, any sequence $(\rho_n)_{n\geq 1}$ of n-player equilibria Fatou converges in probability to ρ.

Proof. If (i) holds, then $\rho^m(t+) = \rho^M(t)$ for all $t \geq 0$ by Corollary 2.4.4, and (ii) follows since $\rho^m(t+) = \rho^m(t)$ except at the (countably many) jumps of ρ^m. The converse holds because equilibria are right-continuous; cf. Proposition 2.4.1. Finally, if $u + F_t(r - cu) = 1$ has a unique solution, this solution is necessarily increasing-transversal since $u + F_t(r - cu) < 1$ for $u < 0$ and $u + F_t(r - cu) > 1$ for $u > 1$. □

While we will see below that the transversality condition in Theorem 2.5.4 cannot be dropped, we can argue that this condition holds for a generic choice of signals Y^i. More generally, we discuss the following hypothesis (again, note that the extremal solutions can never be decreasing-transversal).

Definition 2.5.6. We say that Hypothesis (H) holds if for all t in a dense subset of $\mathbb{R}_+$, any solution of $u \in [0,1]$ of $u + F_t(r - cu) = 1$ is increasing-transversal or decreasing-transversal.

While this hypothesis does not hold for all choices of Y^i, the exceptional set is small in the sense that a "typical" F_t will not have a local extremum of $u \mapsto u + F_t(r - cu)$ at a solution of $u + F_t(r - cu) = 1$, so that the latter must be transversal. As t varies over $\mathbb{R}_+$, the non-transversal case is somewhat more likely to occur, but typically at only finitely many t so that the hypothesis still holds. There seems to be no obvious way to quantify this. However, we state the following result which confirms the general intuition and shows that Hypothesis (H) is always valid after a small perturbation of Y^i.

Proposition 2.5.7. *For every $\delta > 0$ there exists $0 \leq \varepsilon \leq \delta$ such that after replacing Y_t^i with $Y_t^i + \varepsilon$, Hypothesis (H) is satisfied.*

Proof. Let us first observe that for any real function $f(x)$, the set of local minimum values $S = \{f(x) : x \text{ is a local minimum of } f\}$ is countable. Indeed, for every $s \in S$ there is an open interval I_s with rational endpoints such that $s = \min\{f(x) : x \in I_s\}$. If $s, t \in S$ and $I_s = I_t$, then $s = t$, showing that $I : S \to \mathbb{Q} \times \mathbb{Q}$ is injective.

For fixed $t \geq 0$, denote by $S(t)$ the set of all local minimum and maximum values of $u \mapsto u + F_t(r - cu) - 1$, then $\cup_{t \in \mathbb{Q}} S(t)$ is again countable. Thus, we can find a sequence $a_k \downarrow 0$ with $a_k \notin \cup_{t \in \mathbb{Q}_+} S(t)$. Set $\varepsilon_k = c a_k$. Then, passing from Y_t to $Y_t^{\varepsilon_k} = Y_t + \varepsilon_k$, the function under consideration is

$$u \mapsto u + F_t^{\varepsilon_k}(r - cu) = u + F_t(r - cu - \varepsilon_k) = (u + a_k) + F_t(r - c(u + a_k)) - a_k.$$

By the construction of a_k, we know that 1 is not a local extremum value of this function. However, if a solution of $u + F_t^{\varepsilon_k}(r - cu) = 1$ failed to be transversal, then 1 would be the value at a local extremum. $\square$

2.5.1 Counterexamples

In this section, we illustrate that the assertion of Theorem 2.5.4 may fail without the transversality condition, and more generally that the intervals in Lemma 2.5.2 cannot be improved. The examples presented here are essentially static, meaning that Y_t^i does not depend on t. For purely technical reasons, namely to ensure the finiteness of the optimal stopping times (2.2.3) as assumed throughout, we introduce a time horizon $T \in (0, \infty)$ at which Y_t^i jumps to a value larger than r, thus ensuring that all players stop.

In the first example, we allow for atoms in the distribution of Y_t^i to obtain an analytically tractable example. We argue below that the atoms are not essential to the observed phenomenon.

Example 2.5.8. Let $r = c = 1$ and let $Y_t^i = Y_0^i$, $0 \leq t < T$ be constant i.i.d. processes such that $\mathrm{Law}(Y_t^i) = \frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_2$ for all $0 \leq t < T$, and set $Y_t^i = 2$ for $t \geq T$. Then the law of the minimal n-player equilibrium $\rho_n^m(t)$ converges to $\frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_1$ for all $0 \leq t < T$.

Proof. Proposition 2.3.1 yields two cases for every ω. If strictly less than $n/2$ of the realizations $\{Y_0^i(\omega), i = 1, \ldots, n\}$ equal 2, all players i with $Y_0^i(\omega) = 2$ stop at $t = 0$ and those with $Y_0^i(\omega) = 1/2$ never stop. Whereas if $n/2$ or more of the realizations equal 2, then all agents stop at $t = 0$. It follows that the law of $\rho_n^m(t) \equiv \rho_n^m(0)$ converges to $\frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_1$ as $n \to \infty$. $\square$

The limit law $\frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_1$ can be seen as a mixture of the deterministic mean field equilibria $\rho^m(t) \equiv \frac{1}{2}$ and $\rho^{mrt}(t) \equiv 1$. In fact, with an appropriate definition allowing for randomized equilibria, this mixture is itself an equilibrium. However, a remarkable conclusion is that there are no n-player equilibria converging to the minimal equilibrium ρ^m.

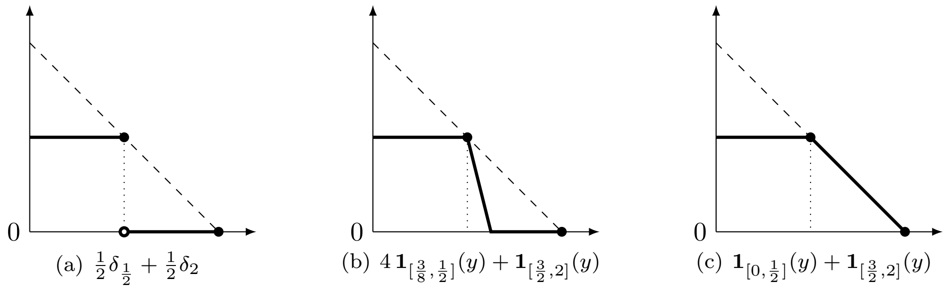


Figure 2.3: Graphs of $F_t(1-u)$ (solid) and $1-u$ (dashed)

Corollary 2.5.9. *In the context of Example 2.5.8, $\rho^m(t)$ is not a weak accumulation point of n-player equilibria, for any $0 \le t < T$.*

Proof. Suppose that there exists a subsequence $\rho_k = \rho_{n_k}$ of n_k-player equilibria such that $\rho_k(t) \to \rho(t) = 1/2$ weakly. Then $\rho_k(t) \ge \rho_k^m(t)$ and $\mathrm{Law}(\rho_k^m(t)) \to \frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_1$ yield a contradiction. ◻

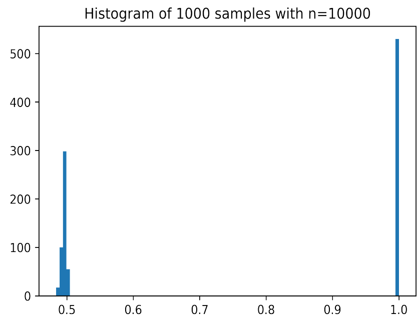
It may be useful to contrast this with the fact that ρ^m is a limit of *approximate* Nash equilibria. To wit, if all players i with $Y_0^i(\omega) = 2$ stop at $t = 0$ whereas those with $Y_0^i(\omega) = 1/2$ do not stop until T, we obtain an approximate Nash equilibrium converging to ρ^m as $n \to \infty$.

The following example is a smooth version of Example 2.5.8 where Y_t^i admits a density; see also Figure 2.3(b). It is not analytically tractable but the qualitative behavior is the same.

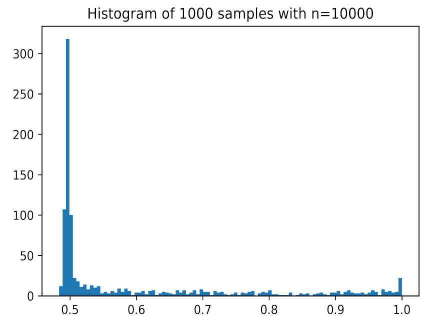
Example 2.5.10. Let $r = c = 1$ and let $Y_t^i = Y_0^i$, $0 \le t < T$ be i.i.d. processes such that the law of Y_t^i has the density $f_t(y) = 4\,\mathbf{1}_{[\frac{3}{8},\frac{1}{2}]}(y) + \mathbf{1}_{[\frac{3}{2},2]}(y)$ for all $0 \le t < T$, and let $Y_t^i = 2 + X^i$, $t \ge T$, where X^i are i.i.d. with a continuous distribution on $[0,1]$. Then the simulation of $\rho_n^m(t)$, cf. Figure 2.4(a), shows that $\rho_n^m(t)$ again converges to $\frac{1}{2}\delta_{\frac{1}{2}} + \frac{1}{2}\delta_1$ for $0 \le t < T$ which is again a mixture of the deterministic mean field equilibria $\rho^m(t) \equiv \frac{1}{2}$ and $\rho^{mrt}(t) \equiv 1$.

In the third example, the mean field game admits a continuum of solutions; see also Figure 2.3(c).

Example 2.5.11. Consider the setting of Example 2.5.10 with density $f_t(y) = \mathbf{1}_{[0,\frac{1}{2}]}(y) + \mathbf{1}_{[\frac{3}{2},2]}(y)$. In this case, we again have $\rho^m(t) \equiv \frac{1}{2}$ and $\rho^{mrt}(t) \equiv 1$, but now all values in between also correspond to mean field equilibria. The simulation of $\rho^m_n(t)$, cf. Figure 2.4(b), illustrates that the law of $\rho^m_n(t)$ converges to a mixture of all these equilibria.



(a) density $4\,\mathbf{1}_{[\frac{3}{8},\frac{1}{2}]}(y) + \mathbf{1}_{[\frac{3}{2},2]}(y)$



(b) density $\mathbf{1}_{[0,\frac{1}{2}]}(y) + \mathbf{1}_{[\frac{3}{2},2]}(y)$

Figure 2.4: Simulations for n-player minimal equilibria ($n = 10'000$). Locations k/n of equilibria with k stopped players on the x-axis, number of samples with that equilibrium on the y-axis.

When the minimal mean field equilibrium is not increasing-transversal, the preceding examples illustrate that it need not be the limit of the minimal n-player equilibria. The final example shows that both cases are possible: it may be the limit even if it is not increasing-transversal.

Example 2.5.12. Consider the setting of Example 2.5.10 with density $f_t(y) = 2\,\mathbf{1}_{[1/2,1]}(y)$. In this case, we easily compute that $\rho^m(t) \equiv 0$ and $\rho^{mrt}(t) \equiv 1$. Nevertheless, $\rho^m_n(t) \equiv 0$ due to $Y^i_t < r$ a.s., and thus $\rho^m_n(t) \to \rho^m(t)$.

2.6 Convergence to General Equilibria

Theorem 2.5.4 shows that if the minimal and maximal mean field equilibria are increasing-transversal (on a dense set), then they are the limits of the minimal and maximal n-player equilibria. Indeed, the latter are obvious candidates for sequences converging to these mean field equilibria. For mean field equilibria that are not extremal, there are no obvious candidates for the approximating n-player equilibria. The following result shows that increasing-transversal equilibria are still limits; however, the approximating n-player equilibria have no simple description. We will see in Section 2.6.2 that the analogue for decreasing-transversal solutions fails.

2.6.1 Increasing-Transversal Equilibria

Theorem 2.6.1. *Let ρ be a mean field equilibrium. Suppose that for all t in a dense subset $D \subseteq \mathbb{R}_+$, the solution $u := \rho(t)$ of $u + F_t(r - cu) = 1$ is increasing-transversal. Then there exist n-player equilibria $(\rho_n)_{n\geq 1}$ which Fatou converge in probability to ρ as $n \to \infty$.*

The first step of the proof is to solve a static version of the problem. This will be accomplished by a fixed point argument for monotone functions.

Lemma 2.6.2. *Let $t \geq 0$, let $u \in [0, 1]$ be an increasing-transversal solution of $u + F_t(r - cu) = 1$ and let $\varepsilon, \delta > 0$. There are $n_0 \in \mathbb{N}$ and $A \in \mathcal{G}_t$ with $P(A) > 1 - \varepsilon$ such that for all $n \geq n_0$ and $\omega \in A$, there exists $k(\omega) \in \mathbb{N}$ such that $|u - k(\omega)/n| \leq \delta$ and (2.3.1) holds; i.e.,*

$$\#\{Y_t^i(\omega) + c\,\frac{k(\omega) - 1}{n} \geq r\} = k(\omega) \quad \textit{and} \quad \#\{Y_t^i(\omega) + c\,\frac{k(\omega)}{n} < r\} = n - k(\omega).$$

Moreover, $k(\omega)$ can be chosen as a measurable function of $Y_t^1(\omega), \dots, Y_t^n(\omega)$.

Proof. Since u is increasing-transversal, there are points $u_0, u_1 \in \mathbb{R}$ such that $u - \delta/2 \leq u_0 < u < u_1 \leq u + \delta/2$ and

$$u_0 < 1 - F_t(r - cu_0) \leq 1 - F_t(r - cu_1) < u_1,$$

where the inequality in the middle is due to the monotonicity of F_t. The Glivenko–Cantelli theorem then implies that the event A_n consisting of all ω such that

$$[nu_0] \leq \#\{Y_t^i(\omega) + c\,\frac{[nu_0] - 1}{n} \geq r\} \leq \#\{Y_t^i(\omega) + c\,\frac{[nu_1]}{n} \geq r\} \leq [nu_1]$$

satisfies $P(A_n) \to 1$. For fixed n and $\omega \in A_n$, consider the integer-valued function

$$k \mapsto G(k) := \#\{Y_t^i(\omega) + c\,\tfrac{k}{n} \geq r\}.$$

By the above, G maps $\{[nu_0]-1, [nu_0], \ldots, [nu_1]\}$ into $\{[nu_0], \ldots, [nu_1]\}$. Moreover, G is monotone increasing. Lemma 2.6.3 below then yields the existence of $[nu_0] \leq k \leq [nu_1]$ such that $G(k-1) = G(k) = k$ which is exactly (2.3.1). By the choice of u_0, u_1 we also have $|u - k/n| \leq \delta$ for n large. Moreover, it is clear from the proof of Lemma 2.6.3 that k is a measurable function of $Y_t^1, \ldots, Y_t^n$. $\square$

Lemma 2.6.3. *Let $x_0 < x_1 < \cdots < x_N$ be real numbers for some $N \geq 1$. Let $J = \{x_1, \ldots, x_N\}$ and $J_0 = \{x_0\} \cup J$. If $f : J_0 \to J$ is monotone increasing, there exists $k \in \{1, \ldots, N\}$ such that $f(x_{k-1}) = f(x_k) = x_k$.*

Proof. Since f is monotone and maps J into J, it must have a fixed point in J. We claim that the minimal $k \in \{1, \ldots, N\}$ such that $f(x_k) = x_k$ has the desired property. Indeed, if $k = 1$, monotonicity implies that $f(x_0) = f(x_1)$ and the proof is complete. If $k > 1$, we observe that $f(x_{l-1}) \geq x_l$ for all $1 \leq l \leq k$. Indeed, $f(x_1) \geq x_2$ since x_1 is not a fixed point, but then $f(x_2) \geq x_3$ since x_2 is not a fixed point and f is monotone, and so on. In particular, $f(x_{k-1}) \geq x_k$ and thus $f(x_{k-1}) = f(x_k) = x_k$. $\square$

Proof of Theorem 2.6.1. Fix $N \in \mathbb{N}$ and let $t_1 < \cdots < t_N$ be in D. For n large enough, Lemma 2.6.2 allows us to find sets $A_l \in \mathcal{G}_{t_l}$ with $P(A_l) > 1 - N^{-2}$ and random variables k_l satisfying $|\rho(t_l) - k_l/n| \leq \delta := 1/N$ and (2.3.1) on A_l, for $1 \leq l \leq N$.

Following Remark 2.3.5, we can construct n-player equilibria ρ_n^l such that $\rho_n^l(t_l) = k_l/n$ on A_l. Next, we argue that these ρ_n^l can be chosen such that

$$\rho_n^1(t_1) \leq \cdots \leq \rho_n^m(t_m) \text{ on } A_1 \cap \cdots \cap A_m, \quad 1 \leq m \leq N. \tag{2.6.1}$$

Indeed, we have $\rho(t_l) \leq \rho(t_{l+1})$ by the increase of ρ. If $\rho(t_l) < \rho(t_{l+1})$, then we can ensure $\rho_n^l(t_l) \leq \rho_n^{l+1}(t_{l+1})$ on $A_l \cap A_{l+1}$ simply by choosing $\delta < |\rho(t_l) - \rho(t_{l+1})|/2$ in Lemma 2.6.2. If $\rho(t_l) = \rho(t_{l+1})$, we can observe that if the construction in the proof of Lemma 2.6.2 is executed twice with t_l and t_{l+1}, then by choosing the same parameters u_0, u_1 the corresponding functions f_l and f_{l+1} satisfy $f_l \leq f_{l+1}$ due to the increase of Y^i. This implies that the corresponding minimal fixed points produced by the proof of Lemma 2.6.3 satisfy $\rho_n^l(t_l) \leq \rho_n^{l+1}(t_{l+1})$.

In view of (2.6.1), we can use Remark 2.3.4(iii) to construct from the equilibria $(\rho_n^l)_{1\le l\le N}$ another n-player equilibrium ϱ_n with the property that $\varrho_n(t_l)=\rho_n^l(t_l)$ for all $1\le l\le N$ on $A^N:=\cap_{l=1}^N A_l$.

To summarize, ϱ_n satisfies $|\rho(t_l)-\varrho_n(t_l)|\le 1/N$ for all $1\le l\le N$ on the set A^N which has probability $P(A^N)\ge 1-N^{-1}$. By letting $t_1,\dots,t_N$ exhaust a countable dense subset $D'\subseteq D\subseteq\mathbb{R}_+$ as $N\to\infty$, this shows that there exist n-player equilibria $(\varrho_n)_{n\ge1}$ such that $\varrho_n(t)\to\rho(t)$ in probability for all $t\in D'$ and the proof is complete. $\square$

Remark 2.6.4. The construction leading to Theorem 2.6.1 is pathwise and thus extends beyond deterministic mean field equilibria. For instance, let ρ^1,ρ^2 be such equilibria satisfying the assumption of Theorem 2.6.1, let $\lambda\in[0,1]$ and suppose that the n-player game admits a set $A\in\mathcal{G}_0$ with $P(A)=\lambda$. Then we can apply the construction separately on A and A^c to find n-player equilibria ρ_n converging to the mixture $\lambda\delta_{\rho^1}+(1-\lambda)\delta_{\rho^2}$ on a dense set. In the same vain, convergence to more general mixtures could be analyzed.

2.6.2 Decreasing-Transversal Equilibria

Let us begin with a simulation and then establish that the observations correspond to a general result.

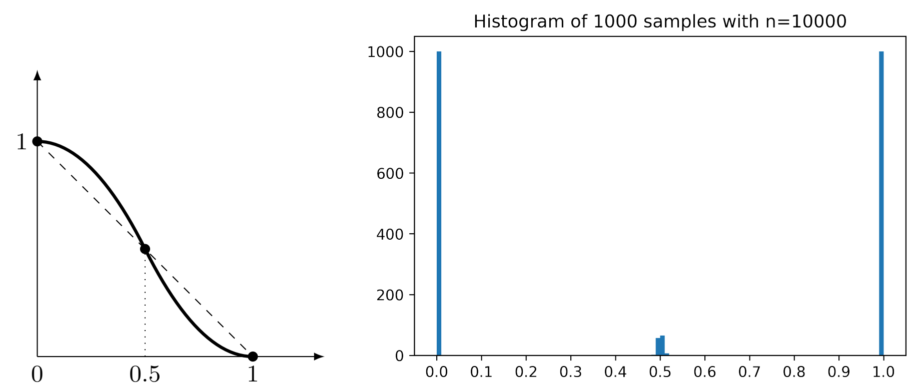


Figure 2.5: C.d.f. and simulation of Example 2.6.5. The decreasing-transversal equilibrium at 0.5 can only be approximated on 12.5% of the samples.

Example 2.6.5. Let $r=c=1$ and let $Y_t^i=Y_0^i$, $0\le t<T$ be constant i.i.d. processes such that $\mathrm{Law}(Y_t^i)$ has the tent-shaped probability density $f(x)=2-4|x-1/2|$, $x\in[0,1]$. As

illustrated in Figure 2.5 (left panel), the corresponding equation (2.4.1) has a decreasing-transversal solution at $u = 1/2$ and increasing-transversal solutions at $u = 0$ and $u = 1$. For the game with $n = 10'000$ players, the histogram in Figure 2.5 shows the values of k/n such that k satisfies the equilibrium conditions (2.3.1). The simulation illustrates the convergence to the equilibria at $u = 0, 1$ as proved in Theorem 2.6.1 but also suggests that $u = 1/2$ is not a limit of n-player equilibria; indeed, only about 12.5% of the samples allow for an n-player equilibrium with k/n close to 1/2. In Proposition 2.6.11, we will establish an asymptotic upper bound which yields $e^{-2} \approx 13.5\%$ in this example.

In the remainder of this section we assume that F_t admits a continuous density f_t. Let $x \in [0, 1]$ be a solution of $u + F_t(r - cu) = 1$. We say that x is *strongly decreasing-transversal* if $\partial_u|_{u=x}[u + F_t(r - cu)] < 0$ or equivalently

$$f_t(r - cx) > c^{-1}.$$

We note that x is then necessarily in $(0, 1)$ and decreasing-transversal in the sense of Definition 2.4.2; the only difference (given the continuity assumption) is that we exclude the case where $u + F_t(r - cu)$ has a vanishing derivative at x (see also Remark 2.6.10). Intuitively, when $f_t(r - cx)$ is large, there are many similar agents (in terms of values of Y^i and relative to the interaction constant c) close to such a state. As a result, these agents may tend to coordinate and either all stop or all not stop: it may be impossible to break up the group[7] and create an n-player equilibrium close to x.

Theorem 2.6.6. *Let ρ be a mean field equilibrium and suppose that the set*

$$\{t \geq 0 : \rho(t) \text{ is strongly decreasing-transversal}\}$$

has nonempty interior.[8] *Then there does not exist a sequence of n-player equilibria ρ_n Fatou converging to ρ in probability.*

[7]Clearly, this intuition does not explain the phase-transition character of the phenomenon. To gather the intuition for a large density, it may be useful to consider the limiting case of an atom in F_t: all agents corresponding to the atom make the same stopping decision.

[8]Note that the condition is nonempty interior rather than the set being nonempty. This corresponds to the fact that convergence in probability on a dense set of times t is sufficient for Fatou convergence; cf. Definition 2.5.3.

This theorem follows from Corollary 2.6.8 below which shows non-existence with positive probability at any fixed time t where $\rho(t)$ is strongly decreasing-transversal. For brevity, we set

$$G_{n,t}(k) = \#\{Y_t^i + c\,\tfrac{k}{n} \geq r\}$$

so that the n-player equilibrium conditions (2.3.1) can be expressed concisely as $G_{n,t}(k) = k = G_{n,t}(k-1)$. Moreover, we introduce

$$\mathcal{K}_{n,t} = \{0 \leq k \leq n : G_{n,t}(k) = k = G_{n,t}(k-1)\}.$$

Roughly speaking, we think of $\mathcal{K}_{n,t}(\omega)$ as the set of all k such that $k/n = \rho_n(t)(\omega)$ for some n-player equilibrium $\rho_n(t)$. (This is not quite meaningful since equilibria can always be altered on nullsets.) More precisely, we have that if ρ_n is a given equilibrium, then $n\rho_n(t) \in \mathcal{K}_{n,t}$ a.s. by (3.1). In particular, we will use below that $\{|x - \rho_n(t)| < \varepsilon\} \subseteq \{\exists\, k \in \mathcal{K}_{n,t} : |x - \frac{k}{n}| < \varepsilon\}$ a.s. for all $x \in [0,1]$ and $\varepsilon > 0$. Finally, we also introduce the superset

$$\mathcal{K}^*_{n,t} = \{0 \leq k \leq n : G_{n,t}(k) = k\} \supseteq \mathcal{K}_{n,t}$$

which has no direct interpretation in terms of our game but is conveniently related to crossings of empirical distribution functions (see the proof below).

Proposition 2.6.7. *Fix $t \geq 0$ and let $x \in (0,1)$ satisfy $x + F_t(r - cx) = 1$. Let $\alpha := cf_t(r - cx)$ and assume that $\alpha > 1$. Then*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} P(\exists\, k \in \mathcal{K}^*_{n,t} : |x - \tfrac{k}{n}| < \varepsilon) = \frac{1-\theta}{\alpha - 1} < 1$$

where $\theta \in (0,1)$ is defined through $\theta e^{-\theta} = \alpha e^{-\alpha}$.

Proof. We first observe the local nature of the claim. Indeed, introducing the uniform random variables $U^i = F_t(Y_t^i)$ we see that the event

$$\begin{aligned} A_{n,\varepsilon} &= \{\exists\, k \in \mathcal{K}^*_{n,t} : |x - \tfrac{k}{n}| < \varepsilon\} \\ &= \{\exists\, 0 \leq k \leq n : \#\{Y_t^i + c\,\tfrac{k}{n} \geq r\} = k,\ |x - \tfrac{k}{n}| < \varepsilon\} \\ &= \{\exists\, 0 \leq k \leq n : \#\{U^i \geq F_t(r - c\,\tfrac{k}{n})\} = k,\ |x - \tfrac{k}{n}| < \varepsilon\} \end{aligned}$$

depends only on the values of F_t in an ε-neighborhood of x. In particular, for ε small enough, we may change F_t outside that neighborhood to guarantee that the set of solutions of $u + F_t(r - cu) = 1$ is $\{0, x, 1\}$.

Considering the c.d.f. $G(u) = 1 - F_t(r - cu)$, the proposition can be rephrased as the probability of having no crossings of the empirical distribution of G and the (theoretical) uniform distribution near x:

$$A_{n,\varepsilon} = \{\exists\, t \in [0,1] : \tfrac{1}{n}\#\{G^{-1}(U^i) \leq t\} = t,\ |x - t| < \varepsilon\}.$$

(To see this identity, note that $\frac{1}{n}\#\{G^{-1}(U^i) \leq t\} = t$ implies $t = k/n$ for some $0 \leq k \leq n$.) Following [91], this problem can be related to boundary-crossing probabilities of Poisson processes which turn out to be computable. In particular, after changing F_t as outlined above, the conditions of [91, Theorem 1] are satisfied for G and noting that $\alpha = G'(x)$, this theorem yields the result. □

In view of $\mathcal{K}_{n,t} \subseteq \mathcal{K}^*_{n,t}$, we have the following consequence (see also Figure 2.6).

Corollary 2.6.8. *Fix $t \geq 0$ and let $x \in [0,1]$ satisfy $x + F_t(r - cx) = 1$. If x is strongly decreasing-transversal, then*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} P(\exists\, k \in \mathcal{K}_{n,t} : |x - \tfrac{k}{n}| < \varepsilon) < 1.$$

Remark 2.6.9. One can ask if the non-existence result is related to the convention made in Section 2.3 that players do not consider their own impact on the state process. To address this question, we can drop the first equation in the equilibrium conditions (2.3.1) and keep only the second (which seems uncontroversial); i.e., $\#\{Y_t^i + c\,\frac{k}{n} < r\} = n - k$.

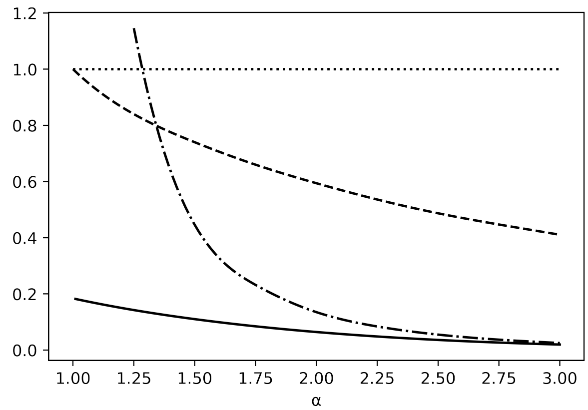


Figure 2.6: Bounds for the probability of finding an n-player equilibrium near x as in Corollary 2.6.8. The dashed and dashed-dotted lines are the upper bounds derived from Proposition 2.6.7 and Proposition 2.6.11, respectively. The solid line is the lower bound from Proposition 2.6.13.

This corresponds to the definition of $\mathcal{K}^*_{n,t}$ and Proposition 2.6.7 shows that non-existence holds even under this condition alone.

Remark 2.6.10. Heuristics suggest that in the tangential case of a decreasing-transversal x with $\alpha = 1$, the limiting probability is 1; i.e., the equilibrium is in fact a limit of n-player equilibria. The tangential case is less important because it generically does not occur, in the same sense as discussed below Definition 2.5.6. We do not provide a rigorous result.

In our last result, we determine the asymptotic expected number of equilibria close to x (for both increasing- and decreasing-transversal cases). Importantly, it implies that this number is positive with positive probability. When $\alpha > 1$ is not close to 1, it also yields a fairly accurate upper bound for the probability of not finding an n-player equilibrium close to x (cf. Example 2.6.5) since the probability of finding more than one solution is small. On the other hand, we see that as $\alpha \to 1$, the expected number of solutions tends to infinity, and in particular the probability of finding many solutions becomes large[9].

Proposition 2.6.11. *Fix $t \geq 0$ and let $x \in (0,1)$ satisfy $x + F_t(r - cx) = 1$. Let $\alpha :=$*

[9]In fact, one can show that $\lim_{\alpha\to 1} \limsup_{n\to\infty} P(\#\{\mathcal{K}_{n,t} \cap |x - \frac{k}{n}| < \varepsilon\} = j) = 0$ for all finite $j \geq 0$ when $\varepsilon > 0$ is small enough.

$cf_t(r - cx)$ and assume that $\alpha \neq 1$. Then

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} E[\#\{k \in \mathcal{K}_{n,t} : |x - \tfrac{k}{n}| < \varepsilon\}] = \frac{e^{-\alpha}}{|1-\alpha|}.$$

In particular,

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} P(\exists\, k \in \mathcal{K}_{n,t} : |x - \tfrac{k}{n}| < \varepsilon) \leq \frac{e^{-\alpha}}{|1-\alpha|}.$$

One consequence of Proposition 2.6.11 is that non-uniqueness is indeed the typical case for the n-player game, as claimed in the Introduction: under the stated smoothness assumption on F_t, we typically have at least one mean field equilibrium corresponding to $0 \neq \alpha < 1$ and then the proposition and Lemma 2.6.2 imply that there is more than one n-player equilibrium, for large n.

Proof of Proposition 2.6.11. We may assume that $c = 1$, and we drop the index t everywhere. We denote

$$\alpha(z) = f(r - z)$$

and recall that $x \in (0,1)$ and $\alpha = \alpha(x) \neq 1$. Fix $\varepsilon > 0$ and denote

$$x_- = x - \varepsilon, \quad x_+ = x + \varepsilon,$$

$$F_- = F(r - x - \varepsilon), \quad F_+ = F(r - x + \varepsilon),$$

$$\alpha_- = \inf_{|z-x|<\varepsilon} \alpha(z), \quad \alpha_+ = \sup_{|z-x|<\varepsilon} \alpha(z),$$

$$m(z) = \inf_{|z-x|\leq\varepsilon} z(1-z), \quad M(z) = \sup_{|z-x|\leq\varepsilon} z(1-z).$$

We assume that ε is small enough such that $x_\pm \in (0,1)$ and $1 \notin [\alpha_-, \alpha_+]$.

Step 1: Bounds for $P(k \in \mathcal{K}_n)$. Fix n and let $U^i = F(Y^i)$, $1 \leq i \leq n$ so that (U^i) are i.i.d. Unif$[0,1]$, and let $U^{(1)} \geq \cdots \geq U^{(n)}$ be the associated reverse order statistics. Noting that $U^{(k)} = U_{(n-k+1)}$ for the usual (increasing) order statistics $U_{(\cdot)}$, we have that $U^{(k)} \sim \mathrm{Beta}(n-k+1, k)$ and $U^{(k+1)} = U^{(k)} W_k^{\frac{1}{n-k}}$ where $W_k \sim$ Unif$[0,1]$ is independent;

cf. [5, Section 4]. Moreover, we note that $k \in \mathcal{K}_n$ is equivalent to

$$U^{(k)} \geq F(r - \tfrac{k-1}{n}) =: F_{k-1} \quad \text{and} \quad U^{(k+1)} \leq F(r - \tfrac{k}{n}) =: F_k. \tag{2.6.2}$$

As a result, for any deterministic integer $1 \leq k \leq n$,

$$\begin{aligned}
P(k \in \mathcal{K}_n) &= P\big(U^{(k+1)} \leq F_k,\, U^{(k)} \geq F_{k-1}\big) \\
&= \int_{F_{k-1}}^{1} P\big(U^{(k+1)} \leq F_k \big| U^{(k)} = z\big)\, dP(U^{(k)} = z) \\
&= \int_{F_{k-1}}^{1} P\big(W \leq (F_k/z)^{n-k} \big| U^{(k)} = z\big)\, dP(U^{(k)} = z) \\
&= \frac{n!}{(n-k)!(k-1)!} \int_{F_{k-1}}^{1} F_k^{n-k} (1-z)^{k-1}\, dz \\
&= \binom{n}{k} F_k^{n-k} (1 - F_{k-1})^k
\end{aligned} \tag{2.6.3}$$

where $dP(U^{(k)} = z)$ indicates integration with respect to the law of $U^{(k)}$. We may observe that this quantity is reminiscent of a binomial distribution except that the success probability changes with k. Next, we use Taylor's theorem to find that

$$F_{k-1} = F(r - \tfrac{k}{n} + \tfrac{1}{n}) = F(r - \tfrac{k}{n}) + \alpha_k/n = F_k + \alpha_k/n \tag{2.6.4}$$

where $\alpha_k = \alpha(\eta_k)$ with $\eta_k \in [\frac{k-1}{n}, \frac{k}{n}]$ and in particular $\alpha_k \in [\alpha_-, \alpha_+]$. Now suppose that $|x - \frac{k}{n}| < \varepsilon$. Then $k \geq nx_-$, and using also $F_k \geq F_-$,

$$\begin{aligned}
P(k \in \mathcal{K}_n) &= \binom{n}{k} F_k^{n-k} (1 - F_k - \alpha_k/n)^k \\
&= \binom{n}{k} F_k^{n-k} (1 - F_k)^k \left(1 - \frac{\alpha_k}{(1-F_k)n}\right)^k \\
&\leq \binom{n}{k} F_k^{n-k} (1 - F_k)^k \left(1 - \frac{\alpha_-}{(1-F_-)n}\right)^{nx_-}.
\end{aligned}$$

The fact that

$$(1-y) \leq e^{-y} \leq (1-y)(1+o(y))$$

as $y \to 0$ applied with $y = w/n$ yields

$$(1 - \tfrac{w}{n})^n \leq e^{-w} \leq (1 - \tfrac{w}{n})^n \, (1 + O(1/n))$$

as $n \to \infty$, uniformly over w in a compact interval. This leads us to the upper bound

$$P(k \in \mathcal{K}_n) \leq \binom{n}{k} F_k^{n-k} (1 - F_k)^k e^{-\frac{\alpha_- x_-}{1 - F_-}}. \tag{2.6.5}$$

Similarly, we have the lower bound

$$\begin{aligned} P(k \in \mathcal{K}_n) &\geq \binom{n}{k} F_k^{n-k} (1 - F_k)^k \left(1 - \frac{\alpha_+}{(1 - F_+)n}\right)^{nx_+} \\ &\geq \binom{n}{k} F_k^{n-k} (1 - F_k)^k e^{-\frac{\alpha_+ x_+}{1 - F_+}} \, (1 + O(1/n)). \end{aligned}$$

Step 2: Decay away from x. Let us recall Robbin's version [98] of the Stirling approximation,

$$\sqrt{2\pi n} (\tfrac{n}{e})^n e^{\frac{1}{12n+1}} \leq n! \leq \sqrt{2\pi n} (\tfrac{n}{e})^n e^{\frac{1}{12n}}, \tag{2.6.6}$$

showing in particular that $n! = \sqrt{2\pi n}\, (\frac{n}{e})^n \, (1 + O(1/n))$. Since $n - k$ and k are comparable to n when $|x - \frac{k}{n}| < \varepsilon$, we have

$$\binom{n}{k} = (1 + O(1/n)) \frac{\sqrt{2\pi n}\, (\frac{n}{e})^n}{\sqrt{2\pi (n-k)}\, (\frac{n-k}{e})^{(n-k)} \sqrt{2\pi k}\, (\frac{k}{e})^k}$$

uniformly over all k such that $|x - \frac{k}{n}| < \varepsilon$. This shows that

$$\begin{aligned} Z_{n,\varepsilon} &:= \sum_{k\colon |x - \frac{k}{n}| < \varepsilon} \binom{n}{k} F_k^{n-k} (1 - F_k)^k \\ &= (1 + O(1/n)) \sum_{k\colon |x - \frac{k}{n}| < \varepsilon} \frac{1}{\sqrt{2\pi n (1 - \frac{k}{n}) \frac{k}{n}}} \frac{F_k^{n-k} (1 - F_k)^k}{(1 - \frac{k}{n})^{n-k} (\frac{k}{n})^k} \\ &\leq \frac{1 + O(1/n)}{\sqrt{m(x)}} \sum_{k\colon |x - \frac{k}{n}| < \varepsilon} \frac{1}{\sqrt{2\pi n}} \frac{F_k^{n-k} (1 - F_k)^k}{(1 - \frac{k}{n})^{n-k} (\frac{k}{n})^k}. \end{aligned} \tag{2.6.7}$$

Our next goal is to estimate the summand above. We introduce the function

$$\varphi(z) = (1-z)^{n-k} z^k$$

so that

$$\frac{F_k^{n-k}(1-F_k)^k}{(1-\frac{k}{n})^{n-k}(\frac{k}{n})^k} = \frac{\varphi(1-F_k)}{\varphi(\frac{k}{n})} \tag{2.6.8}$$

is the term in question. We can use Taylor's theorem similarly as above to find

$$F_k = F(r - \tfrac{k}{n}) = F(r - x + x - \tfrac{k}{n}) = F(r-x) + \tilde{\alpha}_k(x - \tfrac{k}{n})$$

where $\tilde{\alpha}_k \in [\alpha_-, \alpha_+]$. As $F(r-x) = 1-x$, this equality can be rewritten as

$$F_k = 1 - \tfrac{k}{n} + (\tilde{\alpha}_k - 1)(x - \tfrac{k}{n}).$$

Introducing also

$$\psi(z) = \log\varphi(z) = (n-k)\log(1-z) + k\log z,$$

we have

$$\psi'(z) = -\frac{n-k}{1-z} + \frac{k}{z}, \quad \psi''(z) = -n\left[\frac{1-\frac{k}{n}}{(1-z)^2} + \frac{\frac{k}{n}}{z^2}\right] < 0$$

and then $\psi'(k/n) = 0$ shows that ψ and φ have a global maximum at k/n. Taylor's theorem at the second order yields

$$\psi(1-F_k) - \psi(\tfrac{k}{n}) = \psi(\tfrac{k}{n} - (\tilde{\alpha}_k - 1)(x - \tfrac{k}{n})) - \psi(\tfrac{k}{n}) = \frac{\psi''(\xi_k)}{2}(\tilde{\alpha}_k - 1)^2(x - \tfrac{k}{n})^2$$

for a suitable number ξ_k between $\frac{k}{n}$ and $\frac{k}{n} - (\tilde{\alpha}_k - 1)(x - \frac{k}{n})$. Therefore, we have $|\xi_k - x| < A\varepsilon$, with $A = \max\{\alpha_+, 1\}$. Using the above formula for $\psi''(z)$ and setting

$$\Gamma_\varepsilon = \inf_{\substack{|p-x|<\varepsilon \\ |z-x|<A\varepsilon}} \left[\frac{1-p}{(1-z)^2} + \frac{p}{z^2}\right],$$

we arrive at

$$\psi(1-F_k) - \psi(\tfrac{k}{n}) \le -\frac{n}{2}\Gamma_\varepsilon(\tilde{\alpha}_k - 1)^2(x - \tfrac{k}{n})^2.$$

Setting also $\alpha_* = \alpha_-$ if $\alpha > 1$ and $\alpha_* = \alpha_+$ if $\alpha < 1$, exponentiating leads us to the desired estimate

$$\frac{\varphi(1-F_k)}{\varphi(\frac{k}{n})} \leq \exp\left(-\frac{n}{2}\Gamma_\varepsilon(\alpha_*-1)^2(x-\tfrac{k}{n})^2\right)$$

and plugging this into (2.6.7) we have that

$$Z_{n,\varepsilon} \leq \frac{1+O(1/n)}{\sqrt{m(x)}} \sum_{k\colon |x-\frac{k}{n}|<\varepsilon} \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{n}{2}\Gamma_\varepsilon(\alpha_*-1)^2(x-\tfrac{k}{n})^2\right).$$

Set

$$w_k = \sqrt{n\Gamma_\varepsilon}|\alpha_*-1|(\tfrac{k}{n}-x)$$

and note that

$$\Delta w := w_k - w_{k-1} = \frac{1}{\sqrt{n}}\sqrt{\Gamma_\varepsilon}|\alpha_*-1|.$$

The above sum can then be written as

$$\begin{aligned} Z_{n,\varepsilon} &\leq \frac{1+O(1/n)}{\sqrt{m(x)}} \sum_{k\colon |w_k|<\sqrt{n\Gamma_\varepsilon}|\alpha_*-1|\varepsilon} \frac{1}{\sqrt{2\pi n}}\, e^{-w_k^2/2} \\ &= \frac{1+O(1/n)}{\sqrt{m(x)\Gamma_\varepsilon}} \frac{1}{|\alpha_*-1|} \sum_{k\colon |w_k|<\sqrt{n\Gamma_\varepsilon}|\alpha_*-1|\varepsilon} \frac{1}{\sqrt{2\pi}}\, e^{-w_k^2/2}\Delta w \end{aligned}$$

which suggests comparison with a Gaussian integral $\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\, e^{-w^2/2}\, dw = 1$. Indeed, after subtracting the two largest summands neighboring the origin, the sum can be seen as a Riemann sum which is entirely below the integral. These two summands are $O(1/\sqrt{n})$ so that

$$\sum_{k\colon |w_k|<\sqrt{n\Gamma_\varepsilon}|\alpha_*-1|\varepsilon} \frac{1}{\sqrt{2\pi}}\, e^{-w_k^2/2}\Delta w \leq 1 + O(1/\sqrt{n})$$

and finally

$$Z_{n,\varepsilon} \leq \frac{1}{\sqrt{m(x)\Gamma_\varepsilon}} \frac{1}{|\alpha_*-1|}\,(1+O(1/\sqrt{n})).$$

Step 3: Conclusion. Recalling (2.6.5) we have

$$
\begin{aligned}
E[\#\{k \in \mathcal{K}_n : |x - \tfrac{k}{n}| < \varepsilon\}] &= \sum_{k:\, |x-\frac{k}{n}|<\varepsilon} P(k \in \mathcal{K}_n) \\
&\le e^{-\frac{\alpha_- x_-}{1-F_-}} Z_{n,\varepsilon} \\
&\le e^{-\frac{\alpha_- x_-}{1-F_-}} \frac{1}{\sqrt{m(x)\Gamma_\varepsilon}} \frac{1}{|\alpha_* - 1|} \left(1 + O(1/\sqrt{n})\right)
\end{aligned}
$$

and hence

$$
\limsup_{n\to\infty} E[\#\{k \in \mathcal{K}_n : |x - \tfrac{k}{n}| < \varepsilon\}] \le e^{-\frac{\alpha_- x_-}{1-F_-}} \frac{1}{\sqrt{m(x)\Gamma_\varepsilon}} \frac{1}{|\alpha_* - 1|}.
$$

As $\varepsilon \to 0$ we have $x_- \to x$, $\alpha_- \to \alpha$, $\alpha_* \to \alpha$, $F_- \to F(r-x) = 1-x$ and

$$
m(x) \to x(1-x), \quad \Gamma_\varepsilon \to \frac{1}{1-x} + \frac{1}{x} = \frac{1}{x(1-x)}.
$$

Thus,

$$
\limsup_{\varepsilon\to 0} \limsup_{n\to\infty} E[\#\{k \in \mathcal{K}_n : |x - \tfrac{k}{n}| < \varepsilon\}] \le \frac{e^{-\alpha}}{|\alpha - 1|}.
$$

The matching lower bound follows similarly after replacing α_- by α_+, F_- by F_+, and so on. $\square$

Remark 2.6.12. The above proof offers insight into the speed of convergence of n-player equilibria. Specifically, the estimates entail that if $\varepsilon_n \downarrow 0$ is such that $\varepsilon_n\sqrt{n} \to \beta \in [0,\infty]$, then

$$
E[\#\{k \in \mathcal{K}_{n,t} : |x - \tfrac{k}{n}| < \varepsilon_n\}] \to \frac{e^{-\alpha}}{|1-\alpha|}\mu\left(-\frac{|\alpha-1|}{\sqrt{x(1-x)}}\beta, \frac{|\alpha-1|}{\sqrt{x(1-x)}}\beta \right)
$$

where μ is the standard Gaussian distribution. Thus, a ball of radius $r_n/\sqrt{n}$ around x, where $r_n \to \infty$ arbitrarily slowly, will asymptotically contain all n-player equilibria converging to x, and this is optimal in the sense that if $\limsup r_n < \infty$ the ball will miss some solutions.

In our final result we complement the upper bound in Proposition 2.6.11 by a lower bound. The gap between the bounds vanishes for large α; see also Figure 2.6.

Proposition 2.6.13. *Fix $t \geq 0$, let $x \in (0,1)$ satisfy $x + F_t(r - cx) = 1$ and suppose that $\alpha := cf_t(r - cx) > 1$. Then*

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} P(\exists\, k \in \mathcal{K}_{n,t} : |x - \tfrac{k}{n}| < \varepsilon) \geq L(\alpha) > 0$$

where

$$L(\alpha) = \frac{e^{-\alpha}}{(\alpha - 1)\left(1 + 2\sqrt{\frac{2}{|a_0|}}\left\{1 - \Phi\left(\sqrt{2|a_0|}\right)\right\}\right)}$$

with $a_0 := 1 - \alpha + \log(\alpha) < 0$ and Φ is the standard normal c.d.f.

Since the lower bound is strictly positive, we can interpret the result as stating that x is necessarily part of a mixture which is itself a limit of n-player equilibria. In summary, when x is strongly decreasing-transversal, we cannot find n-player equilibria converging to x at time t, but at least we can find n-player equilibria converging to a randomized mean field equilibrium which charges x.

Proof of Proposition 2.6.13. We use the notation from the proof of Proposition 2.6.11 and suppress t. Let $\mathcal{K} = \mathcal{K}_{n,t}$ and $X = X_{n,\varepsilon} = \#\{k \in \mathcal{K}_{n,t} : |x - \frac{k}{n}| \leq \varepsilon\}$. Set $\mu = E[X]$ and let

$$A = A_{n,\varepsilon} = \{|X - c\mu| \geq c\mu\}$$

for a constant $c > 0$ to be chosen later. Clearly $P(X = 0) \leq P(A)$. Using the Markov inequality

$$P(|X - c\mu| \geq c\mu) \leq \frac{E\left((X - c\mu)^2\right)}{c^2\mu^2} = \frac{E\left(X^2\right)}{c^2\mu^2} - \frac{2}{c} + 1 = \frac{2}{c}\left(\frac{E\left(X^2\right)}{2c\mu^2} - 1\right) + 1$$

and choosing $c = \theta \frac{E[X^2]}{2\mu^2}$ for some $\theta > 1$, we obtain that

$$P(X = 0) \leq 1 - \frac{4\mu^2(\theta - 1)}{\theta^2 E[X^2]}.$$

Optimizing over the right-hand side, we note that $\theta = 2$ yields the best bound, so we choose

$c = \frac{E[X^2]}{\mu^2}$ and conclude that

$$P(X=0) \leq 1 - \frac{\mu^2}{E[X^2]} = 1 - \frac{E[X]^2}{E[X^2]}. \tag{2.6.9}$$

Since we have already determined the limit of $E[X]$ in Proposition 2.6.11, our goal is to find an upper bound for $E[X^2]$. To that end, we first compute

$$P(k \in \mathcal{K}, j \in \mathcal{K}) = P(U^{(k+1)} \leq F_k, U^{(k)} \geq F_{k-1}, U^{(j+1)} \leq F_j, U^{(j)} \geq F_{j-1})$$

for $k < j$; recall the notation of (2.6.2). In fact, this probability is zero for $j = k+1$, so we focus on $k+2 \leq j$. Conditionally on $U^{(k+1)} = h < U^{(k)} = u$, the pair $(U^{(j)}, U^{(j+1)})$ has the same distribution as $(hV^{(j-(k+1))}, hV^{(j-k)})$ where $V^{(\ell)}$ are the reverse order statistics of an i.i.d. sample $V_1, \cdots, V_{n-(k+1)}$ of size $n-(k+1)$ and distribution Unif$[0,1]$. Thus, we have

$$\begin{aligned} &P\left(U^{(j+1)} \leq F_j, U^{(j)} \geq F_{j-1} \middle| U^{(k+1)} = h, U^{(k)} = u\right) \\ &= P\left(V^{j-(k+1)} \leq \frac{F_j}{h}, V^{(j-k)} \geq \frac{F_{j-1}}{h}\right). \end{aligned} \tag{2.6.10}$$

Clearly $P(V^{(j-k)} \geq \frac{F_{j-1}}{h}) = 0$ if $F_{j-1} \geq h$, so we only need to consider the case $h \in [F_{j-1}, F_k]$. Using the formula developed in (2.6.3), we obtain

$$\begin{aligned} &P\left(U^{(j+1)} \leq F_j, U^{(j)} \geq F_{j-1} \middle| U^{(k+1)} = h, U^{(k)} = u\right) \\ &= \binom{n-(k+1)}{j-(k+1)} \left(\frac{F_j}{h}\right)^{n-j} \left(1 - \frac{F_{j-1}}{h}\right)^{j-(k+1)}. \end{aligned} \tag{2.6.11}$$

As above (2.6.2), the joint density of $U^{(k)}$ and $U^{(k+1)}$ can be computed using the fact that $U^{(k)} \sim \text{Beta}(n-k+1, k)$ and $U^{(k+1)} = W_k^{\frac{1}{n-k}} U^{(k)}$ where $W_k \sim \text{Unif}[0,1]$ is independent of $U^{(k)}$:

$$dP\left(U^{(k)} = u, U^{(k+1)} = h\right) = k(n-k)\binom{n}{k}(1-u)^{k-1} h^{n-(k+1)} \mathbf{1}_{0 \leq h \leq u \leq 1}\, du\, dh.$$

Integrating with respect to this density and using the appropriate restrictions, we deduce that

$$
\begin{aligned}
P(k \in \mathcal{K}, j \in \mathcal{K}) = k(n-k)\binom{n}{k}\binom{n-(k+1)}{j-(k+1)} F_j^{n-j} \\
\times \int_{F_{k-1}}^{1} (1-u)^{k-1}\, du \int_{F_{j-1}}^{F_k} (h - F_{j-1})^{j-(k+1)}\, dh \\
= \binom{n}{k}(1-F_{k-1})^k \frac{n-k}{j-k}\binom{n-(k+1)}{j-(k+1)} F_j^{n-j}(F_k - F_{j-1})^{j-k} \\
= \binom{n}{k}(1-F_{k-1})^k F_k^{n-k}\binom{n-k}{j-k}\left(\frac{F_j}{F_k}\right)^{n-j}\left(1-\frac{F_{j-1}}{F_k}\right)^{j-k} \\
\leq \binom{n}{k}(1-F_{k-1})^k F_k^{n-k}\binom{n-k}{j-k}\left(\frac{F_j}{F_k}\right)^{n-j}\left(1-\frac{F_j}{F_k}\right)^{j-k}.
\end{aligned}
$$

By a repeated application of (2.6.4) we have that $\frac{F_j}{F_k} = 1 - \frac{\alpha_j(j-k)}{nF_k}$ for some $\alpha_j \in [\alpha_-, \alpha_+]$ and hence the last two terms above satisfy

$$
\begin{aligned}
\left(\frac{F_j}{F_k}\right)^{n-j}\left(1-\frac{F_j}{F_k}\right)^{j-k} &\leq \left[1-\frac{\alpha_j(j-k)}{nF_k}\right]^{n-j}\left[\frac{\alpha_j(j-k)}{nF_k}\right]^{j-k} \\
&\leq \exp\left(-\alpha_-(j-k)\frac{n-j}{nF_k}\right)(\alpha_+)^{j-k}\left(\frac{j-k}{nF_k}\right)^{j-k} \\
&\leq \exp\left(-\alpha_-(j-k)\frac{1-x_+}{F_+}\right)(\alpha_+)^{j-k}\left(\frac{j-k}{nF_k}\right)^{j-k}.
\end{aligned}
$$

On the other hand, Stirling's approximation as in (2.6.6) yields

$$
\begin{aligned}
&\binom{n-k}{j-k}\left(\frac{j-k}{nF_k}\right)^{j-k} = \frac{(n-k)!}{(n-j)!(j-k)!}\left(\frac{j-k}{nF_k}\right)^{j-k} \leq \left(\frac{n-k}{nF_k}\right)^{j-k}\frac{(j-k)^{j-k}}{(j-k)!} \\
&\leq \left(\frac{1-x_-}{F_-}\right)^{j-k}(j-k)^{j-k}\left[\left(\frac{(j-k)}{e}\right)^{j-k}\sqrt{2\pi(j-k)}\exp\left(\frac{1}{12(j-k)+1}\right)\right]^{-1} \\
&\leq \left(\frac{1-x_-}{F_-}\right)^{j-k}\frac{e^{j-k}}{\sqrt{2\pi(j-k)}}.
\end{aligned}
$$

As a result, we obtain the upper bound

$$
P(k \in \mathcal{K}, j \in \mathcal{K}) \leq \binom{n}{k}(1-F_{k-1})^k F_k^{n-k}\frac{1}{\sqrt{2\pi(j-k)}}\exp(a(j-k)) \tag{2.6.12}
$$

where

$$
a = a(\alpha, \varepsilon) := 1 - \alpha_-\frac{1-x_+}{F_+} + \log(\alpha_+) + \log\left(\frac{1-x_-}{F_-}\right).
$$

If the following sums run over indices i with $|x-i/n| \le \varepsilon$, we can express the second moment of X as

$$
\begin{aligned}
E[X^2] &= E\left[\left(\sum\nolimits_k \mathbf{1}_{k\in\mathcal{K}}\right)\left(\sum\nolimits_j \mathbf{1}_{j\in\mathcal{K}}\right)\right] = E\left[\sum\nolimits_{k,j} \mathbf{1}_{k\in\mathcal{K}}\mathbf{1}_{j\in\mathcal{K}}\right] \\
&= E\left[\sum\nolimits_k \mathbf{1}_{k\in\mathcal{K}} + 2\sum\nolimits_{k<j} \mathbf{1}_{k\in\mathcal{K}}\mathbf{1}_{j\in\mathcal{K}}\right] \\
&= \sum\nolimits_k P(k\in\mathcal{K}) + 2\sum\nolimits_{k<j} P(k\in\mathcal{K},\, j\in\mathcal{K}).
\end{aligned}
$$

Thus, (2.6.12) leads to

$$
\begin{aligned}
E[X^2] &= \sum_{k\colon |x-k/n|\le\varepsilon} P(k\in\mathcal{K}) + 2 \sum_{\substack{k,j\colon j\ge k+2,\\ |x-k/n|\le\varepsilon,\\ |x-j/n|\le\varepsilon}} P(k\in\mathcal{K}, j\in\mathcal{K}) \\
&\le E[X] + \frac{2}{\sqrt{2\pi}}\, E[X] \sum_{\ell=2}^{n(x_+-x_-)} \frac{1}{\sqrt{\ell}}\, e^{a\ell}.
\end{aligned}
$$

Note that $a_0 := \lim_{\varepsilon\downarrow 0} a(\alpha,\varepsilon) = 1-\alpha+\log(\alpha)$ is strictly negative since $\alpha > 1$. Thus, $a = a(\alpha,\varepsilon) < 0$ for ε small enough, so that $\frac{1}{\sqrt{\ell}}\, e^{a\ell}$ is summable. More precisely,

$$
\frac{1}{\sqrt{2\pi}} \sum_{\ell=2}^{\infty} \frac{1}{\sqrt{\ell}}\, e^{a\ell} \le \frac{1}{\sqrt{2\pi}} \int_1^\infty \frac{1}{\sqrt{\ell}} e^{a\ell}\, d\ell = \sqrt{\frac{2}{|a|}}\, \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2|a|}}^\infty e^{\frac{-z^2}{2}}\, dz.
$$

Recalling also that $\lim_{\varepsilon\to 0}\lim_{n\to\infty} E[X] = \frac{e^{-\alpha}}{|\alpha-1|} =: H(\alpha)$ by Proposition 2.6.11, we deduce that

$$
\limsup_{\varepsilon\to 0}\limsup_{n\to\infty} E[X^2] \le H(\alpha)\left(1 + 2\sqrt{\frac{2}{|a_0|}}\left(1-\Phi\left(\sqrt{2|a_0|}\right)\right)\right)
$$

and combining this with (2.6.9) yields the claim. $\square$

Chapter 3

Asset Pricing, Heterogeneous Beliefs and Liquidity

3.1 Introduction

Heterogeneous beliefs about fundamental values are a key motive for trade on financial markets. Accordingly, a rich literature studies how prices form as the aggregate of subjective beliefs; see e.g. the survey [101] for numerous references. This synthesis happens by means of trading: agents with lower individual valuations sell to agents who are more optimistic about fundamentals. Hence, *liquidity*—the ease with which trades can be implemented—plays an important role in determining how beliefs are reflected in prices.

In the present study, we propose a tractable model that allows us to study the interplay of heterogenous beliefs and liquidity in determining asset prices. We consider N (types of) agents who have different beliefs about the state process determining the payoff of a given asset. They trade the asset in continuous time to maximize their expected returns, penalized with quadratic costs on inventories and trading rates. We show that this model admits a unique Markovian equilibrium. The equilibrium price is characterized as the solution of a *linear* system of parabolic equations with a weak coupling (i.e., the equations are coupled only through the zeroth-order terms). The solution, as well as the necessary estimates on its derivatives, are obtained by combining a fixed point argument of [8] for reaction–diffusion equations, classical Schauder theory for parabolic equations and a gradient estimate that seems to be novel.

This characterization allows us to study the influence of the two costs. The holding costs on inventories, parametrized by a coefficient γ, can be seen as a proxy for risk aversion, whereas the costs on the trading rate, with coefficient $\lambda > 0$, stand in for the liquidity (or transaction) cost caused by market impact. These costs determine how agents take into account current and future expected returns when choosing their portfolios, and turn out to play inverse roles. Specifically, when the asset is in zero net supply (a natural assumption for derivative contracts, say) the two costs only enter through their ratio γ/λ. For a positive supply, the asset price remains invariant if the inverse of the supply is rescaled in the same manner as transaction and holding costs, so that the larger trading and holding costs of bigger asset positions are offset by reducing both frictions.

Explicit asymptotic formulas obtain in the limiting regimes where either transaction costs or holding costs are small ($\gamma/\lambda \approx \infty$ or $\gamma/\lambda \approx 0$, respectively). For small transaction costs $\lambda \to 0$, a singular perturbation expansion identifies the leading-order correction term relative to the frictionless equilibrium in which assets are priced by taking conditional expectations under a representative agent's probability measure that averages the agents' beliefs. The correction term turns out to be proportional to the square root $\sqrt{\lambda}$ of the transaction costs. The corresponding constant of proportionality is related to the average of the subjective integrated drifts of the agents' frictionless portfolios. Thus, equilibrium prices increase relative to their frictionless counterparts if agents on average expect to increase their positions in the future, and vice versa. The interpretation is that in illiquid markets, agents take into account their future trading needs to reduce transaction costs. Accordingly, expectations of future purchases already lead to increased positions earlier on and equilibrium prices increase according to the excess demand created by the aggregated adjustments of all agents, and vice versa.

The equilibrium for small holding costs $\gamma \to 0$ can be approximated by a regular perturbation expansion around the risk-neutral equilibrium price which averages all agents' subjective conditional expectations. Here, the leading-order correction term is determined by γ times the average of the agents' expectations of their future positions. Other things equal, agents reduce the magnitude of their positions when holding costs are introduced, thereby reducing the demand of agents who expect to be long on average, and increasing the demand of agents that expect to be short. The resulting sign of the price correction therefore depends on the aggregate expectations in the market.

To illustrate the implications of these results and test the numerical accuracy of the expansions, we consider an example where the state process determining the asset's payoff has Ornstein–Uhlenbeck dynamics. Agents agree on the mean-reversion level and volatility, but disagree about the speed of mean-reversion. For these linear state dynamics, the parabolic PDE system describing the equilibrium price can be reduced to a system of linear ODEs by a suitable ansatz. The corresponding equilibrium prices can be computed numerically using standard ODE solvers and compared to the explicit formulas that obtain for our small-cost asymptotics in this case. We find that the introduction of small transaction costs increases volatility, in line with the asymmetric information model of [42], the risk-sharing model studied in [62] and empirical studies such as [57, 71, 109]. By contrast, the introduction of small holding costs decreases the equilibrium volatility. The reason is the opposite manner in which the two costs influence how agents take into account future trading opportunities. Without transaction costs, agents who believe in faster than average mean-reversion perceive a mean-reverting price process and therefore sell when its value is high, whereas agents who believe in slower mean-reversion perceive a price process that exhibits "momentum" and therefore buy in this case. While this frictionless tradeoff only depends on the current dynamics of the asset, transaction costs force the agents to take into account future trading opportunities as well. This makes the current trading opportunities less attractive for the agent believing in faster mean-reversion and therefore creates an excess demand for the asset when its price is high. This in turn further increases high prices and conversely decreases low ones, leading to additional volatility.

Holding costs have the opposite effect, by discounting the importance of future trading opportunities and therefore reducing volatility relative to the risk-neutral limiting price. In fact, numerical experiments suggest that the exact equilibrium volatility smoothly interpolates between the risk-neutral volatility (which is highest) and its counterpart without transaction costs (which is lowest). For model parameters estimated from time series data for the USD/EUR exchange rate, we find that the exact equilibrium prices agree with these comparative statics gleaned from their asymptotic approximations.

For models where trading is frictionless, there is an extensive literature on asset pricing under heterogeneous beliefs; see, e.g., [23, 45, 72, 101] and the references therein. To obtain tractable results with limited liquidity, we focus on a model with quadratic holding and trading costs as well as linear preferences over gains and losses.

Similar linear-quadratic liquidity models are used in partial equilibrium contexts by [1, 4, 6, 79]. Risk-sharing equilibria with homogenous beliefs are studied in [25, 54, 62, 99]. Since the corresponding first-order conditions are linear, these models are considerably more tractable than equilibrium models with other preferences or trading costs, where analytical results are only available if prices or trading strategies are deterministic [87, 110, 111, 112] or agents only trade once [100]. A numerical analysis of an equilibrium model with heterogenous beliefs and transaction costs is carried out in [29].

Considering a holding cost on risky positions as in [36, 40, 94, 99] further simplifies the analysis compared to models where the corresponding risk penalty is imposed on the variance of the risky positions as in [53, 54, 62]. Indeed, in the present model, we can characterize the equilibrium price by a system of linear PDEs, avoiding the nonlinear equations that naturally appear in models where agents have risk aversion in the form of concave utility functions. Since the present work focuses on equilibrium asset prices with heterogeneous *beliefs* about the underlying state process, we abstract from heterogeneous holding costs. These are considered in [62] for agents with homogeneous beliefs and in [37] for a partial equilibrium model with heterogeneous beliefs.

This paper is organized as follows. Section 3.2 details the financial market and the definition of an equilibrium. In Section 3.3 we derive the optimal portfolio of any agent given an exogenous asset price process. Section 3.4 provides the existence, uniqueness and PDE characterization of the equilibrium price. The leading-order asymptotics for small transaction and holding costs are presented in Sections 3.5 and 3.6, respectively. The concluding Section 3.7 covers the example with mean reversion.

Notation. As usual, $C = C(\mathbb{R}^n)$ is the space of continuous functions $g(x)$ on $\mathbb{R}^n$ and C^k is the space of functions $g \in C$ whose partial derivatives up to order k exist and belong to

C. Similarly, $C^{1,k}$ is the space of continuous functions $g(t,x)$ such that $g(t,\cdot), \partial_t g(t,\cdot) \in C^k$. For any of these spaces, a subscript "b" indicates that the functions *and* all mentioned derivatives are bounded. The dimension n of the underlying domain is often understood from the context. Conditional expectations are denoted $E_t[\cdot] = E[\cdot|\mathcal{F}_t]$ for brevity and when F is a functional of the paths of a process Y, we will often write $E_{t,y}[F(Y)] = E[F(Y)|Y_t = y]$. In this context, Y will be the solution of an SDE and $E[F(Y)|Y_t = y]$ can be unambiguously defined as $E[F(Y^{t,y})]$ where $(Y_s^{t,y})_{s\geq t}$ is the unique solution of the corresponding SDE with initial condition y at time t.

3.2 Model

Beliefs. Let X be the coordinate-mapping process on the space $\Omega = C_0([0,T],\mathbb{R}^d)$ of continuous, d-dimensional paths with $\omega_0 = 0$, equipped with the canonical σ-field $\mathcal{F}$ and filtration $(\mathcal{F}_t)$ generated by X. We consider N (types of) agents with heterogeneous views on the distribution of the state process X. Specifically, for each $1 \leq i \leq N$, let Q_i be a probability measure on Ω under which X satisfies

$$dX_t = b_i(t,X_t)dt + \sigma_i(t,X_t)dW_t^i \tag{3.2.1}$$

where W^i is a d'-dimensional Brownian motion. We assume that $b_i : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma_i : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d'}$ are jointly Lipschitz and bounded. This guarantees in particular that (3.2.1) has a unique strong solution. Moreover, we assume that the matrix $\sigma_i^2 := \sigma_i\sigma_i^\top$ is uniformly parabolic: there is a constant $\kappa > 0$ such that $\xi^\top \sigma_i^2 \xi \geq \kappa|\xi|^2$ for all $\xi \in \mathbb{R}^d$. The associated generator is denoted by

$$\mathcal{L}^i = \partial_t + b_i\partial_x + \frac{1}{2}\operatorname{Tr}\sigma_i^2\partial_{xx}. \tag{3.2.2}$$

Remark 3.2.1. The above assumptions imply that the support of Q_i is the whole space Ω; cf. [105, Theorem 3.1]. Thus, if F and G are continuous functions on Ω, then $F(X) = G(X)$ Q_i-a.s. is equivalent to $F = G$. This fact will be used throughout the paper, often implicitly.

Uniform parabolicity is convenient to simplify the exposition, but of course results similar to ours could be obtained under different assumptions. When the support of X is not the whole space, the statements involving price functions and PDEs need to be restricted to a suitable domain (as, e.g., in [90]).

Market Model. Let $f \in C_b^3(\mathbb{R}^d)$. We consider N agents that dynamically trade an asset with a single payoff $f(X_T)$ at the time horizon $T > 0$. Fix a constant $a_0 \geq 0$, the *exogenous supply* at time $t = 0$, and the initial asset allocation $a_i \in \mathbb{R}$ to each agent, where $\sum_{i=1}^N a_i = a_0$. Let $\mathscr{L}^p(Q_i)$ denote the set of progressively measurable processes $\phi = (\phi_t)_{0\leq t\leq T}$ (of appropriate dimension) such that $E^i[\int_0^T \phi_t^p dt] < \infty$. An (admissible) *portfolio* for agent i is a scalar process $\phi \in \mathscr{L}^4(Q_i)$ which satisfies $\phi_0 = a_i$ and is absolutely continuous with rate $\dot\phi \in \mathscr{L}^4(Q_i)$.[1] We say that portfolios ϕ^i, $1 \leq i \leq N$ *clear the market* if

$$\sum_{i=1}^N \phi_t^i = a_0, \quad t \in [0, T]$$

holds pointwise. A *price process* (for f) is a progressively measurable process $S = (S_t)_{0\leq t\leq T}$ which satisfies $S_T = f(X_T)$ and is an Itô process with sufficiently integrable coefficients under each Q_i:

$$dS_t = \mu_t^i dt + \nu_t^i dW_t^i \quad \text{with} \quad \mu^i, \nu^i \in \mathscr{L}^4(Q_i) \tag{3.2.3}$$

for some Q_i-Brownian motion W^i, for all $1 \leq i \leq N$.

Equilibrium. To formulate the agents' optimization criteria, we fix a *holding cost* parameter $\gamma > 0$ and a *transaction cost* parameter $\lambda > 0$. (The boundary cases $\lambda = 0$ and $\gamma = 0$ will be considered in Sections 3.5 and 3.6, respectively.) For a given price process S, agent i maximizes her expected returns, penalized for inventories and trading costs,

$$J^i(\phi) = E^i\left[\int_0^T \left(\phi_t dS_t - \frac{\gamma}{2}\phi_t^2 dt - \frac{\lambda}{2}\dot\phi_t^2 dt\right)\right] \tag{3.2.4}$$

over the set of her admissible portfolios. A portfolio ϕ^i is *optimal* for agent i if it is a maximizer. If S is a price process such that there exist optimal portfolios ϕ^i, $1 \leq i \leq N$

[1] The precise integrability condition is not crucial; we simply need to ensure that the local martingale part of $\int \phi dS$ has vanishing expectation when S is defined as in (3.2.3). In our main equilibrium result the optimal portfolios are bounded.

for the agents which clear the market, then S is an *equilibrium price process.* Finally, $v:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ is an *equilibrium price function* if $v(t,X_t)$ defines an equilibrium price process. We shall be interested in equilibria with prices of this Markovian form; however, the associated portfolios are usually path-dependent in the presence of transaction costs.

3.3 Single-Agent Optimality

As a preparation for the equilibrium result, we first fix agent i and solve her individual optimization problem in the face of an exogenous price process. Similar linear-quadratic optimization problems have been considered, e.g., in [4, 6, 25, 36, 53, 54, 79]; we provide a self-contained proof for the convenience of the reader.

Lemma 3.3.1. *Let $\gamma>0$, $\lambda>0$ and*

$$G(t)=\cosh\left(\sqrt{\frac{\gamma}{\lambda}}(T-t)\right),\quad t\in[0,T]. \tag{3.3.1}$$

Given a price process S as in (3.2.3), the $dt\times Q_i$-a.e. unique optimal portfolio for agent i is

$$\phi_t^i=\frac{G(t)}{G(0)}a_i+\int_0^t\frac{G(t)}{G(s)}E_s^i\left[\int_s^T\frac{G(u)}{G(s)}\frac{\mu_u^i}{\lambda}\,du\right]ds. \tag{3.3.2}$$

In particular, the optimal trading rate is characterized by the random ODE

$$\dot\phi_t^i=\frac{G'(t)}{G(t)}\phi_t^i+E_t^i\left[\int_t^T\frac{G(s)}{G(t)}\frac{\mu_s^i}{\lambda}\,ds\right],\quad \phi_0^i=a_i. \tag{3.3.3}$$

As in the previous literature, the optimal trading strategy tracks a (suitably discounted) average $E_t^i[\int_t^T-\frac{\gamma G(s)}{\lambda G'(t)}\frac{\mu_s^i}{\gamma}\,ds]$ of the future values of the no-transaction cost portfolio μ_t^i/γ obtained by pointwise maximization of the drift of (3.2.4). To wit, illiquidity it accounted for by "aiming in front of the moving target" [53]. Both the tracking speed $-G'(t)/G(t)$ and the discount kernel $K(t,s)=-\gamma G(s)/\lambda G'(t)$ are determined by the ratio γ/λ of holding and transaction costs, with relatively lower transaction costs leading to faster trading and more emphasis on the current returns of the asset.

Proof of Lemma 3.3.1. Direct differentiation shows that $\dot\phi^i$ of (3.3.3) is indeed the derivative of ϕ^i in (3.3.2). Moreover, $\mu^i \in \mathscr{L}^4(Q_i)$ and Doob's inequality imply that $\phi^i \in \mathscr{L}^4(Q_i)$ and then (3.3.3) yields that $\dot{\phi}^i \in \mathscr{L}^4(Q_i)$.

Note that

$$J^i(\phi) = E^i\left[\int_0^T \left(\phi_t \mu_t^i - \frac{\gamma}{2}\phi_t^2 - \frac{\lambda}{2}\dot{\phi}_t^2\right) dt\right]$$

for any portfolio ϕ and that portfolios can be parametrized by their trading rates since the initial allocations are fixed and $\dot{\phi} \in \mathscr{L}^4(Q_i)$ implies $\phi \in \mathscr{L}^4(Q_i)$. The strict concavity of J^i implies that any optimizer is (a.e.) unique and that a trading rate $\dot{\phi}$ is optimal if and only if the Gâteaux derivative $\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[J^i(\dot{\phi}+\varepsilon\dot{\vartheta}) - J^i(\dot{\phi})]$ of (3.2.4) vanishes in all directions $\dot{\vartheta} \in \mathscr{L}^4(Q_i)$; that is,

$$\begin{aligned}
0 &= E^i\left[\int_0^T \left(\mu_t^i \int_0^t \dot{\vartheta}_s ds - \gamma\phi_t^i \int_0^t \dot{\vartheta}_s ds - \lambda\dot{\phi}_t^i\dot{\vartheta}_t\right) dt\right] \\
&= E^i\left[\int_0^T \left(\int_t^T \left(\mu_s^i - \gamma\phi_s^i\right) ds - \lambda\dot{\phi}_t^i\right)\dot{\vartheta}_t dt\right], \quad \dot{\vartheta} \in \mathscr{L}^4(Q_i).
\end{aligned}$$

As $\dot{\vartheta}$ is arbitrary, this is equivalent to $\dot{\phi}_t^i = \frac{1}{\lambda}E_t^i[\int_t^T(\mu_s^i - \gamma\phi_s^i)ds]$, which is in turn equivalent to

$$\dot{\phi}_t^i = M_t^i - \frac{1}{\lambda}\int_0^t (\mu_s^i - \gamma\phi_s^i) ds$$

for some Q_i-martingale M^i. Put differently, $\dot{\phi} \in \mathcal{L}^4(Q_i)$ is optimal if and only if it solves the linear forward-backward SDE

$$d\phi_t = \dot{\phi}_t dt, \quad \phi_0 = a_i, \tag{3.3.4}$$

$$d\dot{\phi}_t = \frac{\gamma}{\lambda}\left(\phi_t - \frac{\mu_t^i}{\gamma}\right) dt + dM_t, \quad \dot{\phi}_T = 0. \tag{3.3.5}$$

Direct computation shows that $(\phi^i, \dot{\phi}^i)$ solves this system: (3.3.4) is trivial and for (3.3.5) we note that

$$
\begin{aligned}
d\dot{\phi}_t^i = & \Bigg\{ \left(\frac{G''(t)}{G(t)} - \frac{G'(t)^2}{G(t)^2} \right) \phi_t^i + \frac{G'(t)}{G(t)} \left(\frac{G'(t)}{G(t)} \phi_t^i + E_t^i \left[\int_t^T \frac{G(s)}{G(t)} \frac{\mu_s^i}{\lambda} \, ds \right] \right) \\
& - \frac{G'(t)}{G(t)^2} E_t^i \left[\int_t^T G(s) \frac{\mu_s^i}{\lambda} \, ds \right] - \frac{\mu_t^i}{\lambda} \Bigg\} dt \\
& + \frac{1}{G(t)} dE_t^i \left[\int_0^T \frac{G(s)\mu_s^i}{\lambda} \, ds \right] = \frac{\gamma}{\lambda} \left(\phi_t^i - \frac{\mu_t^i}{\gamma} \right) dt + dM_t
\end{aligned}
$$

for the Q_i-martingale $M = \int_0^\cdot \frac{1}{G(t)} dE_t^i [\int_0^T \frac{G(s)\mu_s^i}{\lambda} \, ds]$, where $G''(t) = \frac{\gamma}{\lambda} G(t)$ was used. As $G'(T) = 0$, the terminal condition $\dot{\phi}_T^i = 0$ is also satisfied. □

3.4 Equilibrium

The following result establishes the existence and uniqueness of an equilibrium price function $v \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ and its characterization through a weakly-coupled system of linear parabolic equations. Recall that the function G was defined in (3.3.1) as $G(t) = \cosh(\sqrt{(\gamma/\lambda)}(T-t))$.

Theorem 3.4.1. *Let $\gamma > 0$, $\lambda > 0$. The parabolic system*

$$\partial_t v_i + \frac{1}{2} \operatorname{Tr} \sigma_i^2 \partial_{xx} v_i + b_i \partial_x v_i + \frac{G'(t)}{G(t)} (v_i - v) = 0, \quad 1 \le i \le N, \tag{3.4.1}$$

$$v := \frac{1}{N} \sum_{i=1}^N v_i + \frac{\lambda G'(t)}{N G(t)} a_0, \tag{3.4.2}$$

$$v_i(T, \cdot) = f, \quad 1 \le i \le N \tag{3.4.3}$$

has a unique solution $v_1, \dots, v_N \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$, and the function v defined via (3.4.2) is an equilibrium price function. It is unique in the sense that any equilibrium price function $w \in C^{1,2}([0,T] \times \mathbb{R}^d)$ with polynomial growth must be equal to v. The equilibrium portfolios are given by

$$\phi_t^i = \frac{G(t)}{G(0)} a_i + \int_0^t \frac{G(t)}{G(s)} E_s^i \left[\int_s^T \frac{G(u)}{G(s)} \frac{\mathcal{L}^i v(u, X_u)}{\lambda} \, du \right] ds.$$

An immediate consequence of this result is that the holding costs γ and transaction costs λ have dual roles. In particular, in the case of zero net supply $a_0 = 0$, the equilibrium price depends only on the ratio γ/λ, so that small transaction costs are equivalent to large holding costs. When $a_0 > 0$, the theorem shows that the equilibrium price is 0-homogeneous in $(\gamma, \lambda, 1/a_0)$. We discuss this in more detail in Section 3.6 below, after deriving the limiting cases for small costs.

As a preparation for the proof of Theorem 3.4.1, we first establish the analytic properties of the parabolic system. Given $\alpha \in (0,1)$, the Hölder space $C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d)$ consists of the functions $w(t,x)$ such that $w, \partial_t w, \partial_x w, \partial_{xx} w$ exist, are bounded, and uniformly Hölder continuous with exponents $\alpha/2$ in t and α in x.

Proposition 3.4.2. *The system (3.4.1–3.4.3) has a unique solution $v_1, \dots, v_N \in C_b^{1,2}([0,T]\times\mathbb{R}^d)$. In fact, $v_1, \dots, v_N \in C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d)$ for all $\alpha \in (0,1)$ and uniqueness holds in the larger class of functions $w_1, \dots, w_N \in C^{1,2}([0,T)\times\mathbb{R}^d) \cap C([0,T]\times\mathbb{R}^d)$ satisfying $|w_i(t,x)| \le c_1 \exp(c_2|x|^2)$ for some constants $c_1, c_2 \ge 0$.*

Proof. The system (3.4.1–3.4.3) is a weakly coupled, uniformly parabolic linear system; see [51, Chapter 9] for background. Uniqueness in the stated class is a special case of [22, Theorem 1]. An existence result for such linear systems is contained, e.g., in [51, Theorem 3, p. 256], but this does not yield growth estimates of the type we require here. Our system is also covered by a literature on reaction–diffusion systems. Specifically, [8, Theorem 2.4] yields that (3.4.1–3.4.3) has a unique solution $v_1, \dots, v_N \in C^{1,2}([0,T)\times\mathbb{R}^d) \cap C_b([0,T]\times\mathbb{R}^d)$. The main point (which we have not found provided in the literature) is to prove a useful growth estimate on the derivatives, and for that, the key element is to provide a Hölder estimate for v_i.

(i) In this step we show that v_i is globally Lipschitz in x, uniformly in t. Writing $u = (u_1, \dots, u_N)$, our system is of the general form

$$\mathcal{L}^i u_i(t,x) + h(t, u(t,x)) = 0, \quad u_i(T,x) = f_i(x), \quad 1 \le i \le N \tag{3.4.4}$$

satisfying the following conditions, for some constant $c > 0$: the function h is jointly Lipschitz with norm $\mathrm{Lip}(h) \le c$ (hence $h(t,\cdot)$ is of linear growth, uniformly in t); the coefficients of $\mathcal{L}^i$ are bounded and Lipschitz; each f_i is bounded and Lipschitz with norm $\mathrm{Lip}(f_i)$. According to [8, Theorem 2.4], such a system has a (unique) solution $v_1, \dots, v_N \in C^{1,2} \cap C_b$.

Indeed, define

$$F_i(u)(t,x) := E^i\left[f_i(X_T^{t,x}) + \int_t^T h(s,u(s,X_s^{t,x}))ds\right], \quad 1 \le i \le N.$$

It is shown in the proofs of [8, Theorems 2.3 and 2.4] that $F = (F_1,\dots,F_N)$ is a contraction on $(C_b)^N = C_b\times\cdots\times C_b$ for a complete norm which is equivalent to the uniform norm. More precisely, this holds after suitably truncating h (i.e., so that the truncation does not affect the bounded solution). It is shown that if we start at any $u \in (C_b)^N$ and iterate F, the sequence $u^n = (F\circ\cdots\circ F)(u)$ will converge uniformly to a solution $(v_1,\dots,v_N) \in (C^{1,2}\cap C_b)^N$ of (3.4.4). We may, in particular, pick $u \in (C_b)^N$ such that $\sup_{0\le s\le T}\operatorname{Lip}(u_i(s,\cdot)) < \infty$ for all $1 \le i \le N$ as our starting point for the iteration.

By a standard estimate on SDEs (e.g., [107, Theorem 2.4 (i), p. 8]),

$$E^i|X_s^{t,x} - X_s^{t,y}| \le K|x-y|, \quad 0 \le s \le T$$

for a constant K depending only on the Lipschitz constants of the coefficients of $\mathcal{L}^i$ and T. Fix a small time interval $[t,T]$ of length $\tau = T-t > 0$ and let

$$L_u = L_u^{(t)} = \max_{1\le i\le N}\sup_{t\le r\le T}\operatorname{Lip}(u_i(r,\cdot)).$$

Then for $t \le s \le T$ we have that

$$\begin{aligned}
&|F_i(u)(s,x) - F_i(u)(s,y)| \\
&\le E^i\left[\operatorname{Lip}(f_i)|X_T^{s,x} - X_T^{s,y}| + \int_s^T c\max_i\operatorname{Lip}(u_i(r,\cdot))|X_r^{s,x} - X_r^{s,y}|dr\right] \\
&\le [K\operatorname{Lip}(f_i) + \tau cKL_u]\,|x-y|.
\end{aligned}$$

This holds for all i. Choose τ such that $\varepsilon := \tau cK < 1$ and set $L_f = \max_{1\le i\le N} K\operatorname{Lip}(f_i)$, then

$$\operatorname{Lip}(F(u)(s,\cdot)) \le L_f + \varepsilon L_u, \quad t \le s \le T,$$

the notation of course meaning that each component $F_i(u)$ satisfies this property. Iterating yields that $u^n = (F \circ \cdots \circ F)(u)$ satisfies the geometric estimate

$$\mathrm{Lip}(u^n(s,\cdot)) \le L_f \sum_{k=0}^{n-1} \varepsilon^k + \varepsilon^n L_u$$

and hence the uniform limit $(v_1, \dots, v_N) = \lim u^n$ satisfies $\mathrm{Lip}(v_i(s,\cdot)) \le L_f(1-\varepsilon)^{-1}$ for $t \le s \le T$.

Note that the size τ of the interval in the above argument does not depend on $\mathrm{Lip}(f)$. Hence we can repeat the argument on the interval $[T-2\tau, T-\tau]$, replacing the terminal condition f_i by $\tilde{f}_i := v_i(T-\tau,\cdot)$. Continuing finitely many times, we conclude that $\sup_{0\le s\le T} \mathrm{Lip}(v_i(s,\cdot)) < \infty$.

(ii) Next, we show that v_i is globally 1/2-Hölder in t, uniformly in x. A simple SDE estimate shows that

$$E^i|X_s^{t',x} - X_s^{t,x}| \le K|t'-t|^{1/2}, \quad 0 \le t \le t' \le s \le T$$

where K now depends on the Lipschitz constants and uniform bounds for b_i and σ_i as well as T. (To see this one may, e.g., go through the proof of [107, Theorem 2.4 (ii), p. 8] and use the uniform bounds in the estimate below Equation (2.5) of that reference to avoid a dependence on x in the final estimate for $E^i|X_s^{t,x} - X_s^{t,y}|$.)

As mentioned above, the relevant function h in (3.4.4) is truncated in u, so that $\|h\|_\infty < \infty$. This yields the (crude but simple) estimate

$$\int_t^{t'} |h(s, X_s^{t,x}, u(s, X_s^{t,x}))|ds \le \|h\|_\infty |t'-t| \le c'|t'-t|^{1/2} \tag{3.4.5}$$

for some constant c', since $|t'-t| \le T$. If $u \in (C_b)^N$ is Lipschitz in x with constant L_u uniformly in t, we then have, similarly as in (i) but using also (3.4.5),

$$\begin{aligned}
&|F_i(u)(t',x) - F_i(u)(t,x)| \\
&\le E^i \left[\mathrm{Lip}(f_i)|X_T^{t',x} - X_T^{t,x}| + c'|t'-t|^{1/2} + \int_{t'}^T cL_u|X_s^{t',x} - X_s^{t,x}|ds \right] \\
&\le [K\,\mathrm{Lip}(f_i) + c' + TcL_uK]\,|t'-t|^{1/2}.
\end{aligned}$$

Again, we iterate the mapping F to generate $u^n = (F \circ \cdots \circ F)(u)$. By (i) we have that $\sup_n L_{u^n} < \infty$. Hence, the above shows that $|u_i^n(t', x) - u_i^n(t, x)| \le c''|t' - t|^{1/2}$ for a uniform constant c'', and then the same holds for the limit $(v_1, \dots, v_N) = \lim u^n$.

(iii) We have shown above that v is globally Lipschitz in x and $1/2$-Hölder in t; in particular $v_j \in C^{\alpha/2,\alpha}$ for all $\alpha \in (0, 1)$. (See [81, p. 117] for a detailed definition of the Hölder spaces.) For fixed i, we can see v_i as the solution of a *scalar* PDE which contains $(v_j)_j$ as coefficients: $\varphi = v_i$ is the solution of

$$\tilde{\mathcal{L}}\varphi(t, x) + g(t, x) = 0, \quad \varphi(T, \cdot) = f$$

on $[0, T] \times \mathbb{R}^d$ with terminal value $f \in C^{2+\alpha}$, parabolic operator $\tilde{\mathcal{L}}u := \mathcal{L}^i u - u$ and inhomogeneous term $g \in C^{\alpha/2,\alpha}$ defined by

$$g = v_i + \frac{G'(t)}{G(t)} \left(v_i - \frac{1}{N} \sum_{i=1}^{N} v_i + \frac{\lambda G'(t)}{NG(t)} a_0 \right)$$

using the fixed functions $(v_j)_{1 \le j \le N}$. We can now apply a suitable version of the Schauder estimates to conclude that $v_i = \varphi \in C^{1+\alpha/2,2+\alpha}([0, T] \times \mathbb{R}^d)$; cf. [81, Theorem 9.2.3, p. 140]. □

Remark 3.4.3. Suppose that $b_i, \sigma_i, f \in C_b^\infty$ for $1 \le i \le N$. Then we also have $v_i \in C_b^\infty$.

Proof. If $b_i, \sigma_i \in C^\infty([0, T) \times \mathbb{R}^d)$, interior regularity for parabolic systems as stated in [51, Theorem 11, p. 265] immediately yields that the solution from Proposition 3.4.2 is in $C^\infty([0, T) \times \mathbb{R}^d)$. We need to show that the partial derivatives of any order are bounded.

Fix $1 \le i \le N$ and $1 \le k \le d$ and consider the function $\varphi = \partial_{x_k} v_i$. We can differentiate the system (3.4.1) with respect to x_k and rearrange the terms to find that φ is the solution of a scalar parabolic equation

$$\mathcal{L}\varphi(t, x) + g(t, x) = 0, \quad \varphi(T, \cdot) = \partial_{x_k} f$$

on $[0, T] \times \mathbb{R}^d$ with terminal value $\partial_{x_k} f \in C_b^\infty \subseteq C^{2+\alpha}$. Here the inhomogeneity g incorporates all other terms resulting from the differentiated equation: it is a linear combination, with coefficients in $C^\infty([0, T) \times \mathbb{R}^d)$, of the functions v_j, $1 \le j \le N$ as well as their first and second-order spatial derivatives. As $v_j \in C^{1+\alpha/2,2+\alpha}$ by Proposition 3.4.2, we see in particular that $g \in C^{\alpha/2,\alpha}$. Thus, we can conclude from [81, Theorem 9.2.3, p. 140] that $\partial_{x_k} v_i = \varphi \in C^{1+\alpha/2,2+\alpha}([0,T] \times \mathbb{R}^d)$. In particular, the third-order spatial derivatives of v_i are bounded and uniformly Hölder continuous. Moreover, by the parabolic form of the above equation, the same follows for $\partial_t \partial_{x_k} v_i = \partial_t \varphi$.

This argument can be iterated to the higher-order derivatives. □

Proof of Theorem 3.4.1. The formula for the equilibrium portfolios is a direct consequence of Proposition 3.5.1, so we focus on the price.

(i) Let $v_1, \dots, v_N \in C_b^{1,2}$ be the solution from Proposition 3.4.2 and define v by (3.4.2); we show that v is an equilibrium price function. Itô's formula shows that $S_t = v(t, X_t)$ is a price process as defined in (3.2.3); the coefficients μ_i and ν_i are even bounded. In view of (3.4.2), the function $w_i(t,x) := G(t)v_i(t,x)$ satisfies

$$\mathcal{L}^i w_i = G\mathcal{L}^i v_i + G' v_i = G' v$$

and $w_i(T,x) = G(T)v_i(T,x) = f(x)$. Thus, Itô's formula and the boundedness of $\partial_x v_i$ imply that under Q_i we have the Feynman–Kac representation

$$w_i(t,x) = E^i_{t,x}[f(X_T)] - \int_t^T G'(u) E^i_{t,x}[v(u, X_u)]du.$$

As a result,

$$v_i(t,x) = \frac{E^i_{t,x}[f(X_T)]}{G(t)} - \int_t^T \frac{G'(u)}{G(t)} E^i_{t,x}[v(u, X_u)]du \tag{3.4.6}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^d$. Lemma 3.3.1 shows that given the price $S_t = v(t, X_t)$, the portfolio

$$\phi^i_t = \int_0^t \frac{G(t)}{G(s)} E^i_{s,X_s}\Big[\int_s^T \frac{G(u)}{G(s)} \frac{1}{\lambda} \mathcal{L}^i v(u, X_u)\, du\Big] ds + \frac{G(t)a_i}{G(0)} \tag{3.4.7}$$

is optimal for agent i. It remains to prove that these portfolios clear the market. Recalling that $G(T) = 1$, taking expectations in the integration-by-parts formula $\int_s^T G(u)dS_u =$

$G(T)S_T - G(s)S_s - \int_s^T G'(u)S_u du$ and applying (3.4.6) yield

$$
\begin{aligned}
&E^i_{s,x}\left[\int_s^T \frac{G(u)}{G(s)}\frac{1}{\lambda}\mathcal{L}^i v(u,X_u)\,du\right] \\
&= \frac{1}{\lambda G(s)}\left(E^i_{s,x}[f(X_T)] - G(s)v(s,x) - \int_s^T G'(u)E^i_{s,x}[v(u,X_u)]du\right) \\
&= \frac{1}{\lambda}[v_i(s,x) - v(s,x)]. \qquad (3.4.8)
\end{aligned}
$$

In view of (3.4.2), we deduce that

$$
\sum_{i=1}^N E^i_{s,x}\left[\int_s^T \frac{G(u)}{G(s)}\frac{1}{\lambda}\mathcal{L}^i v(u,X_u)\,du\right] = -\frac{G'(s)}{G(s)}a_0.
$$

Using this in (3.4.7) and integrating $-\frac{G'(s)}{G^2(s)} = \partial_s G(s)^{-1}$, we conclude that

$$
\sum_{i=1}^N \phi^i_t = -a_0\int_0^t \frac{G(t)}{G(s)}\frac{G'(s)}{G(s)}\,ds + \frac{G(t)a_0}{G(0)} = a_0
$$

as desired.

(ii) Let $S_t = w(t,X_t)$ be an equilibrium price process for some function $w \in C^{1,2}([0,T]\times\mathbb{R}^d)$ of polynomial growth (or, more generally, $w \in C^{1,2}([0,T)\times\mathbb{R}^d)$ of polynomial growth and locally Hölder continuous on $[0,T]\times\mathbb{R}^d$). We have $w(T,\cdot) = f$ by Remark 3.2.1. Recall that the coefficients $\mu^i_t = \mathcal{L}^i w(t,X_t)$ and $\nu^i_t = \partial_x w(t,X_t)^\top \sigma^i_t$ of (3.2.3) are in $\mathcal{L}^4(Q_i)$ as part of our definition of a price process. We define w_i by the Feynman–Kac formula (3.4.6) with w instead of v. In view of the assumptions on b_i and σ_i, the function w_i has polynomial growth like w. Moreover, by a careful application of standard PDE results, $w_i \in C^{1,2}$ and w_i is a solution of the associated linear PDE (3.4.1). Specifically, we can use an approximation with bounded domains as detailed in [58, Theorem 1, Condition (A3'), Lemma 2 and the comments above it] under the stated conditions on w.

It remains to show (3.4.2). As a consequence of Lemma 3.3.1, the agent's equilibrium portfolios ϕ^i are given by (3.4.7). Since these portfolios clear the market, $\sum_i \phi^i_s = a_0$ and thus $\partial_s \sum_i \frac{\phi_s^i}{G(s)} = a_0 \partial_s G(s)^{-1}$. Reversing the integration by parts (3.4.8), we conclude that

$$\begin{aligned}\sum_{i=1}^{N} \frac{1}{G(s)\lambda}[w_i(s,x) - w(s,x)] &= \sum_{i=1}^{N} \frac{1}{G(s)} E^i_{s,x}\left[\int_s^T \frac{G(u)}{G(s)} \frac{1}{\lambda} \mathcal{L}^i w(u, X_u)\, du\right] \\ &= \partial_s \sum_{i=1}^{N} \frac{\phi_s^i}{G(s)} = a_0 \partial_s G(s)^{-1} = -a_0 \frac{G'(s)}{G(s)^2}\end{aligned}$$

which is equivalent to (3.4.2). We have thus established that $w_1, \dots, w_N \in C^{1,2}$ are a solution of (3.4.1) with polynomial growth. The claim now follows by the uniqueness of the solution as stated in Proposition 3.4.2. □

Remark 3.4.4. The restriction to Markovian equilibria in Theorem 3.4.1 (meaning that the price is a function of t and x) is related to our choice of proof through PDE arguments rather than fundamental. For instance, if the volatility σ_i is the same for all agents, similar arguments could be carried out using Backward SDEs as in [77]. In that framework, one would obtain uniqueness within a class of possibly non-Markovian equilibria and one could also cover beliefs where (3.2.1) is replaced by coefficients that may depend on the path of X.

3.5 Asymptotics for Small Transaction Costs

In this section we provide intuition for the equilibrium price from Theorem 3.4.1 by describing its asymptotics for small transaction costs $\lambda \to 0$. For later comparison, we first consider the model without transaction costs; i.e., $\lambda = 0$. In this case we drop the requirement of absolute continuity in the definition of the admissible portfolios and we also do not enforce the initial holdings a_i (in any event, agents can instantaneously adjust their position after $t = 0$ without incurring costs). The following result, which is a special case of [94, Theorem 2.1 and Remark 3.5], shows that the corresponding equilibrium corresponds to the price of a representative agent with a view $\bar{Q}$ defined by the averaged drift and volatility coefficients.

Proposition 3.5.1. *Let $\lambda = 0$ and $\gamma > 0$. There exists a unique equilibrium price function $v^0 \in C_b^{1,2}$, given by*

$$v^0(t,x) = \bar{E}_{t,x}[f(X_T)] - \frac{(T-t)\gamma a_0}{N}$$

where $\bar{E}$ is the expectation for the probability $\bar{Q}$ under which

$$dX_t = \bar{b}(t, X_t)dt + \bar{\sigma}(t, X_t)dW_t, \quad \bar{b} = \tfrac{1}{N}\textstyle\sum_{i=1}^N b_i, \quad \bar{\sigma}^2 = \tfrac{1}{N}\textstyle\sum_{i=1}^N \sigma_i^2$$

for a Brownian motion W. Equivalently, v^0 is the unique bounded classical solution of

$$\partial_t v + \frac{1}{2}\operatorname{Tr}\bar{\sigma}^2\partial_{xx}v + \bar{b}\partial_x v - \frac{\gamma a_0}{N} = 0, \quad v(T, \cdot) = f. \tag{3.5.1}$$

In equilibrium, the $dt \times Q_i$-a.e. unique optimal portfolio for agent i is

$$\phi_t^{i,0} = \frac{\mathcal{L}^i v^0(t, X_t)}{\gamma}. \tag{3.5.2}$$

In the remainder of this section we denote the equilibrium price from Theorem 3.4.1 by v^λ to emphasize the dependence on λ. Our goal is to compute its leading-order deviation $\lambda^{-1/2}(v^\lambda - v^0)$ from the frictionless equilibrium price v^0 of Proposition 3.5.1. For simplicity, we focus on the case of a one-dimensional state variable ($d = 1$) with smooth drift and diffusion coefficients and terminal condition: $b_i, \sigma_i, f \in C_b^\infty$ for $1 \leq i \leq N$, and hence $v, v_i \in C_b^\infty$ on the strength of Remark 3.4.3.

Theorem 3.5.2. *For fixed holding costs $\gamma > 0$ and small transaction costs $\lambda \to 0$, the equilibrium price function v^λ from Theorem 3.4.1 has the expansion*

$$v^\lambda(t, x) = v^0(t, x) + \sqrt{\lambda}v^*(t, x) + o(\sqrt{\lambda}) \quad \textit{locally uniformly on } [0, T] \times \mathbb{R}. \tag{3.5.3}$$

Here, v^0 is the frictionless equilibrium price from Proposition 3.5.1 and

$$v^*(t, x) = \frac{\sqrt{\gamma}}{N}\sum_{i=1}^N \bar{E}_{t,x}\left[\int_t^T \mathcal{L}^i\hat{\phi}^{i,0}(s, X_s)ds\right] \tag{3.5.4}$$

where $\hat{\phi}^{i,0}(s, x) = \mathcal{L}^i v^0(s, x)/\gamma$ is the feedback function determining agent i's frictionless optimal portfolio (3.5.2) and the expectation is taken under the probability measure $\bar{Q}$ of the frictionless representative agent for which

$$dX_t = \bar{b}(t, X_t)dt + \bar{\sigma}(t, X_t)dW_t, \quad \bar{b} = \tfrac{1}{N}\textstyle\sum_{i=1}^N b_i, \quad \bar{\sigma}^2 = \tfrac{1}{N}\textstyle\sum_{i=1}^N \sigma_i^2.$$

The singular perturbation expansion (3.5.3) shows that the leading-order deviation of the frictional equilibrium price v^λ from its frictionless counterpart v^0 scales with the square root $\sqrt{\lambda}$ of the trading cost, as in the risk-sharing equilibrium of [62]. With the heterogeneous beliefs considered in the present study, the constant of proportionality (3.5.4) is determined by the integrated drift rates $\int_t^T \mathcal{L}^i \hat{\phi}^{i,0}(s, X_s)ds$ of the agents' frictionless equilibrium portfolios, averaged with respect to agents and states. Thus, equilibrium prices increase relative to their frictionless counterparts if agents on average expect to increase their positions in the future, and vice versa.[2] The interpretation is that in illiquid markets, agents take into account their future trading needs such as to save cumulative transaction costs over the whole time interval. Accordingly, expectations of future purchases already lead to increased positions earlier on, and vice versa. To clear the market, equilibrium prices increase or decrease according to the excess demand or supply created by the aggregated adjustments of all agents.

3.5.1 Proof of Theorem 3.5.2

The first step towards the proof of Theorem 3.5.2 is to show that the functions v_i^λ from Theorem 3.4.1 are not just bounded for each λ, but that this bound is in fact uniform for $\lambda \in (0, \infty)$. In view of the PDEs (3.4.1) from Theorem 3.4.1 and since $\frac{\lambda G'(t)^2}{NG(t)^2} a_0$ is uniformly bounded for all $\lambda > 0$ and $t \in [0, T]$ by the definition of G, this is a special case of the following more general result that will also allow us to derive estimates for small holding costs in the subsequent section.

Lemma 3.5.3. *For $i = 1, \ldots, N$ and an arbitrary parameter $\varepsilon \in \mathscr{E}$, consider functions α_i, β_i, a_i^ε, b_i^ε, $h_i \in C_b^\infty([0, T] \times \mathbb{R})$ and write*

$$\mathcal{L}^i = \partial_t + \frac{1}{2}\beta_i^2 \partial_{xx} + \alpha_i \partial_x.$$

Suppose that $a_i^\varepsilon, b_i^\varepsilon$ are bounded uniformly in $\varepsilon \in \mathscr{E}$ and let $u_i = u_i(\varepsilon, \lambda, \gamma)$, $i = 1, \ldots, N$

[2] Note that while the actual future portfolio changes add up to zero by market clearing, this is not necessarily true for the changes as anticipated by the heterogeneous agents under their subjective probability measures, $\mathcal{L}^i \hat{\phi}^{i,0}$. In the formula for v^*, these anticipated changes are averaged across all states under the probability measure $\bar{Q}$ corresponding to the frictionless representative agent.

denote the unique classical bounded solution of the system

$$\mathcal{L}^i u_i + \Big(\frac{G'}{G} + a_i^\varepsilon\Big) u_i - \frac{G'}{G}\frac{1}{N}\sum_{j=1}^N u_j + b_i^\varepsilon = 0, \quad u_i(T,\cdot) = h_i, \quad i = 1,\dots,N.$$

Then, $|u_i(t,x)| \leq M$ *for a constant* $M > 0$ *independent of* $\varepsilon, \lambda, \gamma \in (0,\infty)$ *and* $(t,x) \in [0,T]\times\mathbb{R}$.

Proof. Existence and uniqueness of the u_i is a special case of [8, Theorem 2.4]. Since these functions are bounded, the Feynman–Kac formula as in [76, Theorem 5.7.6] as well as $G'/G = (\log G)'$ and $G(T) = 1$ give

$$\begin{aligned} e^{\int_0^t a_i^\varepsilon d\tau} u_i(t,x) = E_{t,x}^i \Bigg[&\int_t^T -\frac{G'(s)}{G(t)}\frac{1}{N}\sum_{j=1}^N e^{\int_0^s a_i^\varepsilon d\tau} u_j(s,X_s)ds \\ &+ \int_t^T e^{\int_0^s a_i^\varepsilon d\tau}\frac{G(s)}{G(t)} b_i^\varepsilon ds + e^{\int_0^T a_i^\varepsilon d\tau}\frac{1}{G(t)} h_i(X_T)\Bigg] \end{aligned}$$

where the expectation is taken under the measure for which the state variable has dynamics $dX_t = \alpha_i(t,X_t)dt + \beta_i(t,X_t)dW_t^i$. Choose a uniform bound M for $\big|e^{\int_0^s a_i^\varepsilon d\tau} b_i^\varepsilon\big|$ and $\big|e^{\int_0^T a_i^\varepsilon d\tau} h_i\big|$, and define

$$K(t,\lambda,\gamma) = \max\Big\{ \big|e^{\int_0^t a_i^\varepsilon(\tau,x_\tau)d\tau} u_i(t,x_t)\big| \Big\} < \infty,$$

where the maximum is taken over $i \in \{1,\dots,N\}$, $\varepsilon \in \mathscr{E}$, and $(x_\tau)_{\tau\in[0,t]} \in C_0([0,t],\mathbb{R})$. With this notation,

$$|e^{\int_0^t a_i^\rho d\tau} u_i(t,x)| \leq \int_t^T -\frac{G'(s)}{G(t)} K(s,\lambda,\gamma)ds + M\int_t^T \frac{G(s)}{G(t)}ds + \frac{M}{G(t)},$$

which in turn leads to

$$G(t)K(t,\lambda,\gamma) \leq \int_t^T \Big(-\frac{G'(s)}{G(s)}\Big) G(s)K(s,\lambda,\gamma)ds + M\int_t^T G(s)ds + M.$$

We may read this as an inequality of the form $u(t) \leq \int_t^T B(s)u(s)ds + A(t)$ for $u(t) = G(t)K(t,\lambda,\gamma)$. Using $G'/G = (\log G)'$ and that G is decreasing, $\int_t^T G(r)dr \leq G(t)T$ and

Grönwall's lemma yield

$$G(t)K(t,\lambda,\gamma) \le M(G(t)T+1) - MG(t)\int_t^T \Big(\int_s^T G(r)dr + 1\Big)\frac{G'(s)}{G(s)^2}ds. \tag{3.5.5}$$

Observe that G satisfies $G = \frac{\lambda}{\gamma}G''$ and $G'(T) = 0$ and $\frac{\lambda}{\gamma}\frac{(G')^2}{G^2} \le 1$, so that

$$-\int_t^T \Big(\int_s^T G(r)dr\Big)\frac{G'(s)}{G(s)^2}ds = \int_t^T \frac{\lambda}{\gamma}\frac{G'(s)^2}{G(s)^2}ds \le T - t \le T.$$

Together with

$$-\int_t^T \frac{G'(s)}{G(s)^2}ds = 1 - \frac{1}{G(t)} \le 1,$$

it follows from (3.5.5) and $G(t) \ge 1$ that $K(t,\lambda,\gamma) \le 2M(T+1)$. As a^ε is uniformly bounded in ε, t, x, the function u_i is therefore uniformly bounded in $\varepsilon, \gamma, \lambda, t, x$ by the definition of $K(t,\lambda,\gamma)$. □

Corollary 3.5.4. *There exists $M > 0$ such that $|v_i^\lambda(t,x)| \le M$ for all $\lambda > 0$ and $(t,x) \in [0,T]\times\mathbb{R}$.*

Next, we establish an analogous uniform bound for the derivatives of the functions v_i^λ and v^λ from Theorem 3.5.2.

Lemma 3.5.5. *Fix $k \ge 0$. There exists $M > 0$ such that*

$$|\partial_x^k v_i^\lambda(t,x)|,\ |\partial_x^k v^\lambda(t,x)|,\ |\partial_t\partial_x^k v^\lambda(t,x)| \le M$$

for all $\lambda > 0$ and $(t,x) \in [0,T]\times\mathbb{R}$.

Proof. By Theorem 3.4.1 and Remark 3.4.3, the x-derivatives of the functions v_i^λ, $i = 1,\dots,N$ from Theorem 3.4.1 satisfy the following PDEs obtained by differentiating (3.4.1) with respect to the spatial variable:

$$\begin{aligned}\partial_t\partial_x v_i^\lambda &+ \frac{1}{2}\sigma_i^2\partial_{xx}\partial_x v_i^\lambda + (b_i + \sigma_i\partial_x\sigma_i)\partial_x\partial_x v_i^\lambda \\ &+ \Big(\partial_x b_i + \frac{G'}{G}\Big)\partial_x v_i^\lambda - \frac{G'}{G}\frac{1}{N}\sum_{j=1}^N \partial_x v_j^\lambda = 0, \quad \partial_x v_i^\lambda(T,\cdot) = f'.\end{aligned} \tag{3.5.6}$$

Lemma 3.5.3 therefore yields the desired uniform bound for $|\partial_x v_i^\lambda(t,x)|$, and in turn also for $\partial_x v^\lambda(t,x) = \frac{1}{N}\sum_{i=1}^N \partial_x v_i^\lambda(t,x)$. The corresponding bounds for the higher-order x-derivatives follow by iterating this argument. Finally, the uniform bound for the time derivative of $\partial_x^k v^\lambda$ is then direct consequences of the parabolic form of the PDEs (3.4.1), (3.5.6), etc., and their sums. $\square$

Lemma 3.5.6. *For $\lambda > 0$, consider α, β, a^λ, h of class C_b^∞ and write*

$$\mathcal{L} = \partial_t + \frac{1}{2}\beta^2\partial_{xx} + \alpha\partial_x.$$

Suppose that $w^\lambda \in C_b^\infty$ satisfies $w^\lambda(T,\cdot) = h$ and $\partial_t w^\lambda$, $\partial_x w^\lambda$, $\partial_{xx} w^\lambda$, a^λ are bounded uniformly in λ. Then, the unique bounded classical solution u^λ of

$$\mathcal{L}u^\lambda + a^\lambda(t,x) + \frac{G'(t)}{G(t)}(u^\lambda - w^\lambda) = 0, \quad u^\lambda(T,\cdot) = h$$

satisfies

$$\frac{|u^\lambda(t,x) - w^\lambda(t,x)|}{\sqrt{\lambda}} \le M$$

for some $M > 0$ independent of $\lambda > 0$ and $(t,x) \in [0,T]\times\mathbb{R}$.

Proof. Following the same steps as in the derivation of (3.4.8) yields that

$$u^\lambda(t,x) = w^\lambda(t,x) + E_{t,x}\left[\int_t^T \frac{G(u)}{G(t)}\Big(a^\lambda + \mathcal{L}w^\lambda\Big)(u,X_u)du\right] \tag{3.5.7}$$

where the expectation is taken under the measure for which the state variable has dynamics $dX_t = \alpha(t,X_t)dt + \beta(t,X_t)dW_t$. With a uniform bound M for $a^\lambda + \mathcal{L}w^\lambda$, the desired bound is

$$\frac{|u^\lambda(t,x) - w^\lambda(t,x)|}{\sqrt{\lambda}} \le \frac{M}{\sqrt{\lambda}}\int_t^T \frac{G(u)}{G(t)}du = -\frac{M\sqrt{\lambda}}{\gamma}\frac{G'(t)}{G(t)} \le \frac{M}{\sqrt{\gamma}}$$

where we have once again used $G(u) = \frac{\lambda}{\gamma}G''(u)$ and $G'(T) = 1$ in the second step and the definition of G for the last inequality. $\square$

Corollary 3.5.7. *Fix $k \geq 0$. There exists $M > 0$ such that*

$$\frac{|\partial_x^k v_i^\lambda(t,x) - \partial_x^k v^\lambda(t,x)|}{\sqrt{\lambda}} \leq M$$

for all $\lambda > 0$ and $(t,x) \in [0,T] \times \mathbb{R}$ and $i \in \{1,\ldots,N\}$.

Proof. In view of the PDEs (3.4.1) from Theorem 3.4.1 and the uniform bounds from Lemma 3.5.5, Lemma 3.5.6 yields that $\lambda^{-1/2}|v_i^\lambda - v^\lambda| \leq M$ for some constant M. This proves the assertion for $k = 0$. The analogous bounds for the derivatives follow by applying the same argument to the corresponding PDEs obtained by differentiating (3.4.1) as in the proof of Lemma 3.5.5. $\square$

We can now estimate the difference between v^λ and the frictionless equilibrium price v^0 of Proposition 3.5.1.

Proposition 3.5.8. *Fix $k \geq 0$. There exists $M > 0$ such that*

$$\frac{|\partial_x^k v^\lambda(t,x) - \partial_x^k v^0(t,x)|}{\sqrt{\lambda}}, \; \frac{|\partial_t\partial_x^k v^\lambda(t,x) - \partial_t\partial_x^k v^0(t,x)|}{\sqrt{\lambda}} \leq M$$

for all $\lambda > 0$ and $(t,x) \in [0,T] \times \mathbb{R}$.

Proof. Using (3.4.2) and then (3.4.1) for v, subtracting the PDE (3.5.1) for v^0, and using once again $G(u) = \frac{\lambda}{\gamma}G''(u)$, we obtain

$$\begin{aligned}\partial_t(v^\lambda - v^0) + \frac{1}{2}\bar{\sigma}^2\partial_{xx}(v^\lambda - v^0) + \bar{b}\partial_x(v^\lambda - v^0)& \\ + \frac{1}{N}\sum_{i=1}^N \frac{1}{2}\sigma_i^2(\partial_{xx}v_i^\lambda - \partial_{xx}v^\lambda) + \frac{1}{N}\sum_{i=1}^N b_i(\partial_x v_i^\lambda - \partial_x v^\lambda) &= 0\end{aligned} \tag{3.5.8}$$

with $(v^\lambda - v^0)(T,\cdot) = 0$. Here $\bar{b}, \bar{\sigma}$ are as defined in Proposition 3.5.1. The desired uniform bound for $\lambda^{-1/2}|v^\lambda - v^0|$ is now a consequence of the Feynman–Kac formula and Corollary 3.5.7. The analogous result for $\lambda^{-1/2}|\partial_x^k v^\lambda - \partial_x^k v^0|$ follows from the same argument since Remark 3.4.3 shows that these derivatives satisfy similar PDEs obtained by differentiating (3.5.8). The corresponding bounds for the time derivatives in turn is a direct consequence of the parabolic form of the equations. $\square$

We have the following version of "Laplace's method" for our function $G(t) = \cosh(\sqrt{\frac{\gamma}{\lambda}}(T-t))$ as $\lambda \to 0$.

Lemma 3.5.9. *Given $t \in [0,T)$ and a continuous function F on $[t,T]$,*

$$\sqrt{\frac{\gamma}{\lambda}} \int_t^T \left(\frac{-G'(u)}{G(t)} \left(\int_t^u F(s)ds \right) \right) du \to F(t) \quad \text{as } \lambda \to 0. \tag{3.5.9}$$

Proof. The left-hand side of (3.5.9) can be decomposed as

$$\sqrt{\frac{\gamma}{\lambda}} \int_t^T \frac{-G'(u)}{G(t)} \int_t^u \Big(F(s) - F(t)\Big) ds du + \sqrt{\frac{\gamma}{\lambda}} \int_t^T \frac{-G'(u)}{G(t)} \int_t^u F(t) ds du.$$

Using the uniform continuity of F on $[t,T]$ and observing that $\sqrt{\frac{\gamma}{\lambda}} \frac{-G'(\cdot)}{G(t)}$ converges to 0 locally uniformly on $(t,T]$, one verifies that the first term vanishes for $\lambda \to 0$. Integration by parts and $G = \frac{\lambda}{\gamma} G''$ show that the second term converges to $F(t)$. ☐

Together with the uniform bounds from Proposition 3.5.8, Lemma 3.5.9 allows us to compute the leading-order expansions of $v_i^\lambda - v^\lambda$ and its derivatives.

Lemma 3.5.10. *For $k = 0,1,2$ and $(t,x) \in [0,T) \times \mathbb{R}$, we have*

$$\lim_{\lambda \to 0} \frac{\partial_x^k v_i^\lambda(t,x) - \partial_x^k v^\lambda(t,x)}{\sqrt{\lambda}} = \frac{\partial_x^k \mathcal{L}^i v^0(t,x)}{\sqrt{\gamma}}. \tag{3.5.10}$$

Proof. The proof is similar for $k = 0,1,2$; we only spell it out in the case $k = 2$ for which the computations are most involved. By Theorem 3.4.1 and Remark 3.4.3, the second-order x-derivatives of the functions v_i^λ, $i = 1, \dots, N$ from Theorem 3.4.1 satisfy the following PDEs obtained by differentiating (3.4.1) twice with respect to the spatial variable:

$$\begin{aligned} &\partial_t \partial_{xx} v_i^\lambda + \frac{1}{2}\sigma_i^2 \partial_{xx}\partial_{xx} v_i^\lambda + (b_i + 2\sigma_i \partial_x \sigma_i)\partial_x \partial_{xx} v_i^\lambda \\ &\quad + \Big(c_i + \frac{G'}{G}\Big)\partial_{xx} v_i^\lambda + \partial_{xx} b_i \partial_x v_i^\lambda - \frac{G'}{G}\frac{1}{N}\sum_{j=1}^N \partial_{xx} v_j^\lambda = 0, \quad \partial_{xx} v_i^\lambda(T,\cdot) = f'', \end{aligned}$$

where $c_i = 2\partial_x b_i + (\partial_x \sigma_i)^2 + \sigma_i \partial_{xx}\sigma_i$. Since all functions appearing here are bounded by either by assumption or by Remark 3.4.3, the Feynman–Kac formula and $G'/G = (\log G)'$

yield the stochastic representation

$$
\begin{aligned}
&\partial_{xx} v_i^\lambda(t,x) \\
&= E'_{t,x}\Bigg[\int_t^T \frac{-G'(u)}{G(t)}\Big(e^{\int_t^u c_i d\tau}\partial_{xx}v^\lambda(u,X_u)\Big)du + \frac{e^{\int_t^T c_i d\tau} f''(X_T)}{G(t)} \\
&\qquad + \int_t^T \frac{G(u)}{G(t)}\Big(e^{\int_t^u c_i d\tau}\partial_{xx}b_i\partial_x v_i^\lambda(u,X_u)\Big)du\Bigg]
\end{aligned}
$$

where the expectation $E'[\cdot]$ is taken under the measure Q' for which the state variable has dynamics $dX_t = (b_i + 2\sigma_i\partial_x\sigma_i)(t,X_t)dt + \sigma_i(t,X_t)dW_t^i$. Together with $\int_t^T -\frac{G'(u)}{G(t)}du = 1 - \frac{1}{G(t)}$, this implies

$$
\begin{aligned}
\frac{\partial_{xx}v_i^\lambda - \partial_{xx}v^\lambda}{\sqrt{\lambda}} = E'_{t,x}\Bigg[&\int_t^T \frac{-G'(u)}{\sqrt{\lambda}G(t)}\Big(e^{\int_t^u c_i d\tau}\partial_{xx}v^\lambda(u,X_u) - \partial_{xx}v^\lambda(t,X_t)\Big)du \\
&+ \frac{e^{\int_t^T c_i d\tau} f''(X_T) - \partial_{xx}v^\lambda(t,X_t)}{\sqrt{\lambda}G(t)} \\
&+ \int_t^T \frac{G(u)}{\sqrt{\lambda}G(t)}\Big(e^{\int_t^u c_i d\tau}\partial_{xx}b_i\partial_x v_i^\lambda(u,X_u)\Big)du\Bigg].
\end{aligned} \tag{3.5.11}
$$

Recalling that c_i, f'' and (by Lemma 3.5.5) also $\partial_{xx}v^\lambda$ are bounded (uniformly in λ), dominated convergence and the definition of G show that the expectation of the second term on the right-hand side of (3.5.11) converges to zero as $\lambda \to 0$. Since $\lim_{\lambda\to 0}\frac{G(T)}{\sqrt{\lambda}G(t)} = 0$, dominated convergence and integration by parts show that the expectation of the third term converges to

$$
\begin{aligned}
&E'_{t,x}\left[\lim_{\lambda\to 0}\int_t^T \frac{-G'(u)}{\sqrt{\lambda}G(t)}\Big(\int_t^u e^{\int_t^s c_i d\tau}\partial_{xx}b_i\partial_x v_i^\lambda(s,X_s)ds\Big)du\right] \\
&\quad = \partial_{xx}b_i\partial_x v^0(t,x).
\end{aligned} \tag{3.5.12}
$$

Here we have used Corollary 3.5.7 and Proposition 3.5.8, and Lemma 3.5.9 for the equality. Finally, the expectation of the first term on the right-hand side of (3.5.11) can be rewritten by applying Itô's formula to $e^{\int_t^u c_i d\tau}\partial_{xx}v^\lambda(u,X_u)$, inserting the Q'-dynamics of X and taking into account that the corresponding local martingale part has expectation zero because all involved functions are bounded. Dominated convergence as well as Proposition 3.5.8 and

Lemma 3.5.9 then show that the corresponding limit for $\lambda \to 0$ is

$$E'_{t,x}\left[\lim_{\lambda\to 0}\int_t^T \frac{-G'(u)}{\sqrt{\lambda}G(t)}\int_t^u e^{\int_t^s c_i d\tau}\mathcal{L}_i''\partial_{xx}v^\lambda(s,X_s)dsdu\right] = \mathcal{L}_i''\partial_{xx}v^0(t,x), \qquad (3.5.13)$$

where $\mathcal{L}_i'' = \partial_t + \frac{1}{2}\sigma_i^2\partial_{xx} + (b_i + 2\sigma_i\partial_x\sigma_i)\partial_x + c_i\mathrm{Id}$. The assertion for $k=2$ now follows from (3.5.11–3.5.13) by observing that $\mathcal{L}_i''\partial_{xx} + \partial_{xx}b_i\partial_x = \partial_{xx}\mathcal{L}^i$. $\square$

We can now prove the expansion of the equilibrium price for small transaction costs.

Proof of Theorem 3.5.2. We first observe that the PDE (3.5.8) for $v^\lambda - v^0$ admits the Feynman–Kac representation

$$\begin{aligned}
&(v^\lambda - v^0)(t,x)\\
&= \frac{1}{N}\sum_{i=1}^N \bar{E}_{t,x}\left[\int_t^T \Big(\frac{1}{2}\sigma_i^2(\partial_{xx}v_i^\lambda - \partial_{xx}v^\lambda) + b_i(\partial_x v_i^\lambda - \partial_x v^\lambda)\Big)(s,X_s)ds\right].
\end{aligned}$$

Dominated convergence and the limits computed in Lemma 3.5.10 then yield

$$\begin{aligned}
\lim_{\lambda\to 0}\frac{v^\lambda - v^0}{\sqrt{\lambda}}(t,x) &= \sqrt{\frac{1}{\gamma}}\frac{1}{N}\sum_{i=1}^N \bar{E}_{t,x}\left[\int_t^T \Big(\frac{1}{2}\sigma_i^2\partial_{xx} + b_i\partial_x\Big)\mathcal{L}^i v^0(s,X_s)ds\right]\\
&= \sqrt{\frac{1}{\gamma}}\frac{1}{N}\sum_{i=1}^N \bar{E}_{t,x}\left[\int_t^T \Big(\mathcal{L}^i - \partial_t\Big)\gamma\hat{\phi}^{i,0}(s,X_s)ds\right] \qquad (3.5.14)
\end{aligned}$$

where $\hat{\phi}^{i,0} = \mathcal{L}^i v^0/\gamma$ is the frictionless equilibrium portfolio function of agent i; cf. (3.5.2). Since these strategies clear the market, the sum of their time derivatives is zero and the pointwise limit (3.5.14) simplifies to (3.5.4). The family $\{\lambda^{-1/2}(v^\lambda - v^0)\}_{\lambda>0}$ is bounded and equicontinuous by Proposition 3.5.8; whence, the convergence is in fact locally uniform as a consequence of the Arzelà–Ascoli theorem. $\square$

3.6 Asymptotics for Small Holding Costs

Next, we study the asymptotics of the equilibrium price from Theorem 3.4.1 for small holding costs $\gamma \to 0$ (and fixed transaction costs $\lambda > 0$). To emphasize the dependence on γ, we denote the price v by v^γ in this section. We again focus on the case of a one-dimensional state variable ($d = 1$) with smooth drift and diffusion coefficients and terminal condition; see Remark 3.4.3.

To formulate the result, we first note that the risk-neutral version $\gamma = 0$ of our model is well posed and essentially covered as a simple special case of Theorem 3.4.1 (with the same proof, read with the conventions $G(u)/G(s) = 1$ and $G'(u)/G(s) = 0$). The corresponding equilibrium price is the average of all agents' conditional expectations,

$$v^0(t,x) = \frac{1}{N}\sum_{i=1}^{N} v_i^0(t,x) = \frac{1}{N}\sum_{i=1}^{N} E_{t,x}^i[f(X_T)], \tag{3.6.1}$$

and the corresponding portfolios are

$$\phi_t^{i,0} = a_i + \int_0^t E_s^i\left[\int_s^T \frac{\mathcal{L}^i v^0(u,X_u)}{\lambda}du\right]ds. \tag{3.6.2}$$

(The above notation for the case $\gamma = 0$ should not be confused with the notation for the case $\lambda = 0$ in the preceding section.)

Lemma 3.3.1 shows that when $\gamma > 0$ and $\lambda > 0$, the optimal portfolios take into account future expected returns that are discounted with a kernel determined by γ/λ. As a limiting case, we have seen that the no-transaction-cost portfolio (3.5.2) for $\lambda = 0$ only takes into account the current (subjective) drift rates; this corresponds to an infinite discount. In the opposite extreme, the no-holding cost portfolio (3.6.2) aggregates the future expected returns without discounting.

Accordingly, we expect small holding costs to play a similar role as *large* transaction costs. Indeed, Theorem 3.4.1 shows that when the supply a_0 vanishes, the equilibrium price only depends on the ratio γ/λ—the "urgency parameter" that determines optimal execution trajectories [2] and, more generally, optimal trading strategies with transaction costs in various contexts; cf., e.g., [89] and the references therein. In particular, small holding costs are equivalent to large transaction costs in terms of resulting prices. When $a_0 > 0$, Theorem 3.4.1 shows that the equilibrium price is 0-homogeneous in $(\gamma, \lambda, 1/a_0)$. This means that the asset price remains invariant if the inverse of the supply is rescaled in the same manner as transaction and holding costs, so that the larger trading and holding costs of bigger asset positions are offset by reduced friction coefficients.

The main result of this section is the following regular perturbation expansion for small holding costs $\gamma \to 0$.

Theorem 3.6.1. *For fixed transaction costs $\lambda > 0$ and small holding costs $\gamma \to 0$, the equilibrium price function from Theorem 3.4.1 has the expansion*

$$v^\gamma(t,x) = v^0(t,x) + \gamma v^*(t,x) + o(\gamma) \quad \textit{uniformly on } [0,T] \times \mathbb{R}. \tag{3.6.3}$$

Here v^0 is the equilibrium price (3.6.1) for $\gamma = 0$ and

$$v^*(t,x) = -\frac{1}{N} \sum_{i=1}^{N} E^i_{t,x} \left[\int_t^T \phi_s^{i,0} ds \right]$$

where $\phi_t^{i,0}$ is the optimal strategy (3.6.2) of agent i for $\gamma = 0$ and the expectation is taken under agent i's belief Q_i.

The reference point for the expansion (3.6.3) is the risk-neutral price v^0 of (3.6.1). In this limiting case, agents only consider future expected returns. Other things equal, agents reduce the magnitude of their positions when holding costs are introduced. The above expression for v^* reflects each agent's expectation $E^i_{t,x}[\int_t^T \phi_s^{i,0} ds]$ of their average future position. Adding holding costs reduces the demand by agents who expect to be long on average, and the converse holds for shorts. The resulting sign of the price correction will thus depend on the aggregate expectations in the market. Indeed, the formula for v^* shows that at the first order, the arithmetic average over all agents' expected average positions is the negative of the correction.

Proof. Step 1. Similarly as for Theorem 3.5.2, the first step towards proving this expansion is to establish that the functions v_i^γ from Theorem 3.4.1 are uniformly bounded in γ. Indeed, note that the function $\frac{\lambda G'(t)^2}{NG(t)^2} a_0$ is bounded locally uniformly in γ. Hence, Lemma 3.5.3 applied with the PDEs (3.4.1–3.4.2) from Theorem 3.4.1 yields that given $0 < \bar{\gamma} < \infty$, there exists $M > 0$ such that

$$|v_i^\gamma(t,x)| \leq M \quad \text{for all } \gamma \in [0, \bar{\gamma}]. \tag{3.6.4}$$

Step 2. Next, we show that as $\gamma \to 0$,

$$|v_i^\gamma(t,x) - v_i^0(t,x)| \to 0 \quad \text{uniformly on } [0,T]\times\mathbb{R}. \tag{3.6.5}$$

Indeed, (3.4.1–3.4.2) show that the difference between the functions v_i with and without holding costs satisfies

$$\begin{aligned}\partial_t(v_i^\gamma - v_i^0) &+ \frac{1}{2}\sigma_i^2\partial_{xx}(v_i^\gamma - v_i^0) + b_i\partial_x(v_i^\gamma - v_i^0)\\ &+ \frac{G'(t)}{G(t)}\left(v_i^\gamma - \frac{1}{N}\sum_{j=1}^N v_j^\gamma - \frac{\lambda G'(t)}{NG(t)}a_0\right) = 0, \quad (v_i^\gamma - v_i^0)(T,\cdot) = 0.\end{aligned}$$

Thus, the Feynman–Kac formula yields

$$\begin{aligned}&(v_i^\gamma - v_i^0)(t,x)\\ &\quad= E^i_{t,x}\left[\int_t^T \frac{G'(s)}{G(s)}\Big(v_i^\gamma(s,X_s) - \frac{1}{N}\sum_{j=1}^N v_j^\gamma(s,X_s) - \frac{\lambda G'(s)}{NG(s)}a_0\Big)ds\right]\end{aligned} \tag{3.6.6}$$

where the expectation is taken under agent i's subjective probability measure Q_i. Note that $G'(t)\to 0$ and $G(t)\to 1$ as $\gamma\to 0$, uniformly on $[0,T]$. In view of (3.6.4), we conclude (3.6.5).

Step 3. We can now prove the expansion from Theorem 3.6.1. By (3.6.6),

$$\begin{aligned}&\frac{(v_i^\gamma - v_i^0)(t,x)}{\gamma}\\ &\quad= E^i_{t,x}\left[\int_t^T \frac{G'(s)}{\gamma G(s)}\Big(v_i^\gamma(s,X_s) - \frac{1}{N}\sum_{j=1}^N v_j^\gamma(s,X_s) - \frac{\lambda G'(s)}{NG(s)}a_0\Big)ds\right].\end{aligned}$$

Using the definition of G and Dini's theorem,

$$\lim_{\gamma\to 0}\frac{G'(t)}{G(t)} = 0 \quad \text{and} \quad \lim_{\gamma\to 0}\frac{G'(t)}{\gamma G(t)} = -\frac{T-t}{\lambda}, \quad \text{uniformly on } [0,T]. \tag{3.6.7}$$

Together with (3.6.4), dominated convergence, (3.6.5) and (3.6.1), this yields

$$\lim_{\gamma\to 0}\frac{v_i^\gamma(t,x) - v_i^0(t,x)}{\gamma} = E^i_{t,x}\left[\int_t^T \frac{T-s}{\lambda}\Big(v^0(s,X_s) - v_i^0(s,X_s)\Big)ds\right]$$

uniformly on $[0,T]\times\mathbb{R}$. In view of the definition of v^γ in (3.4.2), and (3.6.7), it follows that

$$\begin{aligned}
&\lim_{\gamma\to 0}\frac{v^\gamma(t,x)-v^0(t,x)}{\gamma}\\
&\quad=\frac{1}{N}\sum_{i=1}^N E^i_{t,x}\left[\int_t^T\frac{T-s}{\lambda}\Big(v^0(s,X_s)-v^0_i(s,X_s)\Big)ds\right]-\frac{(T-t)a_0}{N}.
\end{aligned}\tag{3.6.8}$$

By (3.6.2) and the first identity of (3.4.8) in the special case $\gamma=0$, we have

$$\begin{aligned}
\phi^{i,0}_t-a_i&=\int_0^t E^i_s\left[\frac{1}{\lambda}\Big(f(X_T)-v^0(s,X_s)\Big)\right]ds\\
&=\int_0^t\frac{1}{\lambda}\Big(v^0_i(s,X_s)-v^0(s,X_s)\Big)ds.
\end{aligned}$$

Using this identity to integrate (3.6.8) by parts and taking into account the market-clearing condition $\sum_{i=1}^N\phi^{i,0}=a_0$, the theorem follows. □

3.7 Example: Mean-Reversion Trading

To gain further intuition for the equilibrium of Theorem 3.4.1, we consider an example that can be solved explicitly up to a system of linear ODEs. We will also use this example to test the numerical accuracy of the expansions for small transaction and holding costs relative to the exact solution.

Suppose that $f(x)=x$, so that at time T, the state X represents the asset. Agents believe that X has mean-reverting dynamics

$$dX_t=\kappa_i(\bar{X}-X_t)dt+\sigma dW^i_t.\tag{3.7.1}$$

That is, agents agree on the volatility $\sigma>0$ and the mean-reversion level $\bar{X}>0$, but disagree about the mean-reversion speed $\kappa_i>0$. This can be interpreted as a simple model for a forward contract on a mean-reverting underlying such as an FX rate. As is natural in that context, and to simplify the exposition, we henceforth assume that the net supply of the contract is $a_0=0$.

3.7.1 Equilibrium with Costs

We first consider the exact equilibrium price v with transaction costs $\lambda > 0$ and holding costs $\gamma > 0$ from Theorem 3.4.1. For the linear state dynamics (3.7.1), the parabolic system (3.4.1–3.4.3) can be reduced to a system of linear ODEs by the ansatz

$$v_i^\lambda(t,x) = A_i(t) + B_i(t)x, \quad i = 1, \dots, N.$$

Indeed, writing $\mathbb{1}_N$ and I_N for the $N \times N$-matrices of ones and the identity matrix, respectively, the deterministic functions $B = (B_1, \dots, B_N)^\top$ and $A = (A_1, \dots, A_N)^\top$ satisfy

$$\begin{aligned} B'(t) &= \left[\operatorname{diag}(\kappa_1, \dots, \kappa_n) + \frac{G'(t)}{G(t)}\left(\frac{1}{N}\mathbb{1}_N - I_N\right)\right] B(t), \\ B(T) &= 1. \end{aligned}$$

and

$$\begin{aligned} A'(t) &= \frac{G'(t)}{G(t)}\left(\frac{1}{N}\mathbb{1}_N - I_N\right) A(t) - \bar{X}\operatorname{diag}(\kappa_1, \dots, \kappa_N) B(t), \\ A(T) &= 0. \end{aligned}$$

These ODEs have unique, smooth solutions. Moreover, the equilibrium price then satisfies

$$v(t,x) = \frac{1}{N}\sum_{i=1}^N (A_i(t) + B_i(t)x) = \bar{X} + (x - \bar{X})\frac{1}{N}\sum_{i=1}^N B_i(t), \tag{3.7.2}$$

where we have used the ODEs for the A_i and B_i for the second equality. To be precise, the unbounded terminal conditions and state dynamics (3.7.1) do not satisfy the boundedness assumptions of Theorem 3.4.1. However, with the unique solutions A and B of the above ODEs at hand, the arguments in the proof of Theorem 3.4.1 show that (3.7.2) identifies the unique equilibrium price in the class from smooth functions with linear growth, say.

3.7.2 Transaction-Cost Asymptotics

We first study the equilibrium v^0 with vanishing transactions costs $\lambda = 0$ and fixed holding costs $\gamma > 0$. As the state variable has the dynamics

$$dX_t = \bar{\kappa}(\bar{X} - X_t)dt + \sigma dW_t \quad \text{with} \quad \bar{\kappa} = \frac{1}{N}\sum_{i=1}^{N} \kappa_i \tag{3.7.3}$$

under the aggregate measure $\bar{Q}$, Proposition 3.5.1 yields that

$$v^0(t,x) = \bar{E}_{t,x}\left[X_T\right] - \frac{(T-t)\gamma}{N} a_0 = \bar{X} + (x - \bar{X})e^{-\bar{\kappa}(T-t)}. \tag{3.7.4}$$

As a result, agent i believes that the frictionless equilibrium price has dynamics

$$\begin{aligned} dv^0(t, X_t) &= (\kappa_i - \bar{\kappa})e^{-\bar{\kappa}(T-t)}(\bar{X} - X_t)dt + e^{-\bar{\kappa}(T-t)}\sigma dW_t^i \\ &= (\kappa_i - \bar{\kappa})\big(\bar{X} - v^0(t, X_t)\big)dt + e^{-\bar{\kappa}(T-t)}\sigma dW_t^i. \end{aligned} \tag{3.7.5}$$

This means that agents who believe in faster than average mean-reversion (i.e., $\kappa_i > \bar{\kappa}$) observe a mean-reverting process. By contrast, agents who believe in slower than average mean reversion conclude that the process exhibits "momentum" in that prices above the mean-reversion level are followed by further positive drifts. Whence, in equilibrium, the market is endogenously populated by both "mean-reversion traders" and "trend-followers" even though all agents believe that the underlying has a mean-reverting fundamental value.

Next, we study the leading-order correction $v^\lambda(t,x) - v^0(t,x)$ for $\lambda \to 0$. Again, Theorem 3.5.2 does not apply directly due to the unbounded coefficients, but it is straightforward to carry out the arguments in the proof for the example at hand. Thus, the leading-order correction is $\sqrt{\lambda}v^*(t,x)$ with

$$
\begin{aligned}
v^*(t,x) &= \frac{1}{\sqrt{\gamma}N}\sum_{i=1}^N \bar{E}_{t,x}\left[\int_t^T e^{-\bar{\kappa}(T-s)}(\bar{\kappa}-\kappa_i)^2(X_s-\bar{X})ds\right] \\
&= \frac{1}{\sqrt{\gamma}}\left(\frac{1}{N}\sum_{i=1}^N(\bar{\kappa}-\kappa_i)^2\right)\left[\int_t^T \bar{E}_{t,x}[X_s-\bar{X}]e^{-\bar{\kappa}(T-s)}ds\right] \\
&= \frac{1}{\sqrt{\gamma}}\left(\frac{1}{N}\sum_{i=1}^N(\bar{\kappa}-\kappa_i)^2\right)\left[\int_t^T e^{-\bar{\kappa}(T-t)}(x-\bar{X})ds\right] \\
&= \sqrt{\frac{1}{\gamma}}\left(\frac{1}{N}\sum_{i=1}^N(\bar{\kappa}-\kappa_i)^2\right)(T-t)e^{-\bar{\kappa}(T-t)}(x-\bar{X}). \qquad (3.7.6)
\end{aligned}
$$

Note that $\partial_x v^* \geq 0$, so that the equilibrium volatility is always increased when small transaction costs are added. This is in line with the asymmetric information model of [42], the risk-sharing model studied in [62], and empirical studies such as [57, 71, 109].

In our model, the reason for the increased volatility is that the sign of the correction term v^* is determined by $x-\bar{X}$, so that transaction costs amplify the fluctuations of the frictionless equilibrium price (3.7.4). Let us now discuss why illiquidity affects price levels in this manner. In view of the above formula for v^*, adding small transaction costs increases equilibrium prices when $X_t > \bar{X}$ and reduces prices for $X_t < \bar{X}$. If $X_t > \bar{X}$, agents who believe in larger than average mean-reversion speeds predict the frictionless equilibrium price (3.7.5) to mean-revert downwards towards its long-run mean. Conversely, agents believing in a lower than average mean-reversion speed expect the positive trend to continue and prices to rise even further. Accordingly, the first group of agents wants to sell the asset and the second group wants to purchase it. With small transaction costs added, these trading motives persist, yet changes in portfolios can only be implemented gradually. Accordingly, agents do not only take into account the difference between the current value of the state variable and its long-run mean, but also their expected differences in the future. Since agents believing in faster mean-reversion expect differences to disappear faster, they have a weaker motive to act on the trading opportunities they observe. For $X_t > \bar{X}$, this means that sellers have a weaker motive to trade than buyers, so that prices need to rise in order to clear the market. For $X_t < \bar{X}$, the situation is reversed and small transaction costs decrease prices relative to their frictionless counterparts.

In summary, adding small transaction costs increases prices above the mean-reversion level and decreases price below it, thereby generating larger price fluctuations and a larger equilibrium volatility. By contrast, the *average* price level remains unchanged, in that the correction term mean-reverts around zero under each agent's probability measure. In particular, the simple model considered here—where holding costs are homogenous and long and short positions are treated symmetrically, for example—does not generate the systematic "illiquidity discounts" observed in the empirical literature dating back to [3].

3.7.3 Holding-Cost Asymptotics

We now turn to the small-holding-cost asymptotics from Section 3.6. As a first step, we compute the equilibrium price v^0 with vanishing holding costs $\gamma = 0$ and fixed transactions costs $\gamma > 0$. From (3.6.1), we have

$$v^0(t,x) = \frac{1}{N}\sum_{i=1}^{N} v_i^0(t,x),$$

where

$$v_i^0(t,x) = E^i_{t,x}[X_T] = \bar{X} + (x-\bar{X})e^{-\kappa_i(T-t)}. \tag{3.7.7}$$

By Itô's formula, agent i believes that this risk-neutral equilibrium price has dynamics

$$\begin{aligned}
&dv^0(t,X_t)\\
&= \frac{1}{N}\sum_{j=1}^{N}(\kappa_j-\kappa_i)e^{-\kappa_j(T-t)}(X_t-\bar{X})dt + \frac{1}{N}\sum_{j=1}^{N} e^{-\kappa_j(T-t)}\sigma dW_t^i\\
&= \left(\kappa_i - \frac{\sum_{j=1}^{N}\kappa_j e^{-\kappa_j(T-t)}}{\sum_{j=1}^{N} e^{-\kappa_j(T-t)}}\right)\left(\bar{X} - v^0(t,X_t)\right)dt + \frac{1}{N}\sum_{j=1}^{N} e^{-\kappa_j(T-t)}\sigma dW_t^i.
\end{aligned}$$

The first factor is the difference between κ_i and a (time-dependent) weighted average of $\kappa_1,\ldots,\kappa_N$. Thus, the interpretation is similar as for the equilibrium (3.7.5) with $\lambda = 0$: agents who believe in fast mean reversion observe a mean-reverting asset price whereas agents believing in slow mean reversion perceive momentum. One can also note that the equilibrium volatility without holding costs is always larger than or equal to its counterpart without transaction costs. This follows by applying Jensen's inequality to the gradients of v^0 and (3.7.4).

We now turn to the leading-order correction term for $\gamma \to 0$. Again, the boundedness assumptions in Theorem 3.6.1 are not satisfied in this example, but the arguments in the proof still apply. Accordingly, using the representation (3.6.8), we have

$$
\begin{aligned}
&v^*(t,x) \qquad\qquad (3.7.8)\\
&= \frac{1}{N}\sum_{i=1}^{N} E^i_{t,x}\left[\int_t^T \frac{T-s}{\lambda}\Big(v^0(s,X_s) - v^0_i(s,X_s)\Big)ds\right]\\
&= \frac{1}{N}\sum_{i=1}^{N}\int_t^T \frac{T-s}{\lambda}\Big(\frac{1}{N}\sum_{j=1}^{N} e^{-\kappa_j(T-s)} - e^{-\kappa_i(T-s)}\Big)e^{-\kappa_i(s-t)}(x-\bar{X})ds\\
&= \frac{(x-\bar{X})(T-t)}{\lambda N^2}\left(\sum_{i\neq j}\frac{1}{\kappa_i-\kappa_j}e^{-\kappa_j(T-t)} - \frac{(T-t)(N-1)}{2}\sum_{i=1}^{N} e^{-\kappa_i(T-t)}\right)
\end{aligned}
$$

where the last equality follows from an elementary but lengthy integration.

The Chebychev sum inequality applied to the second representation shows that the coefficient multiplying $x - \bar{X}$ is always negative. Whence, adding small holding costs increases the risk-neutral equilibrium price when the state process X_t is below its mean-reversion level $\bar{X}$ and decreases it when $X_t > \bar{X}$. Since larger holding costs play the same role as lower transaction costs in our model for $a_0 = 0$, the intuition for this is the converse of the argument for adding small transaction costs in Section 3.7.2.

In particular, in view of (3.7.7), adding small holding costs dampens the fluctuations of the risk-neutral equilibrium price and accordingly reduces the equilibrium volatility. This negative effect on the equilibrium volatility and the positive effect of small transaction costs are consistent with the observation made above that the equilibrium volatility without transaction costs always lies below its counterpart with no holding costs. In fact, numerical experiments suggest that the exact equilibrium volatility $\frac{1}{N}\sum_{i=1}^{N} B_i(t)$ from Section 3.7.1 smoothly interpolates between these two extreme cases as γ/λ ranges between ∞ and 0.

3.7.4 A Calibrated Example

To assess the accuracy of the small-cost asymptotics from Sections 3.7.2 and 3.7.3, we now compare the explicit formulas (3.7.6) and (3.7.8) to the numerical solutions of the ODEs from Section 3.7.1 describing the exact equilibrium price. Throughout, we consider a time horizon of $T = 3$ years.

To obtain reasonable values for the other model parameters, we calibrate the state dynamics (3.7.3) to USD/EUR exchange rate data from 2009–2019 available from the website of the St. Louis Fed at https://fred.stlouisfed.org/series/DEXUSEU. The model parameters can then be estimated by matching the first two stationary moments to their empirical counterparts and fitting the (linear) log-autocorrelation function to its empirical counterpart using linear regression. This leads to

$$\sigma = 0.128, \quad \bar{X} = 1.25, \quad \bar{\kappa} = 0.575.$$

With a zero net supply, this suffices to pin down the equilibrium price without transaction costs (3.7.4), since the latter does not depend on the agents' holding costs in this case. For the equilibrium prices with transaction costs (3.7.2), we additionally need to specify each agent's individual belief as well as the transaction cost λ and the holding cost γ. Inspired by similar parameter values used for commodities and equities in [53, 36], respectively, we use

$$\lambda = 10^{-7} \quad \text{and} \quad \gamma = 10^{-8}.$$

The free parameter $\kappa_1 = 2\bar{\kappa} - \kappa_2$ can in turn be chosen arbitrarily to capture the agents' disagreement about the mean-reversion speed of the exchange rate. For

$$\kappa_1 = 3\kappa_2 = 0.8625,$$

and $x = 1$, equilibrium asset prices and volatilities are plotted in Figures 3.1 and 3.2. These numerical examples clearly display the qualitative properties derived from the asymptotic formulas in the previous sections. Indeed, the equilibrium values with both holding and transaction costs always lie between the limiting cases where only one of these costs is present. Moreover, for $X_t = x < \bar{X}$, the equilibrium price with transaction costs lies below its no-transaction cost counterpart and the corresponding volatility is increased by the trading cost, in line with the discussion in Section 3.7.2.

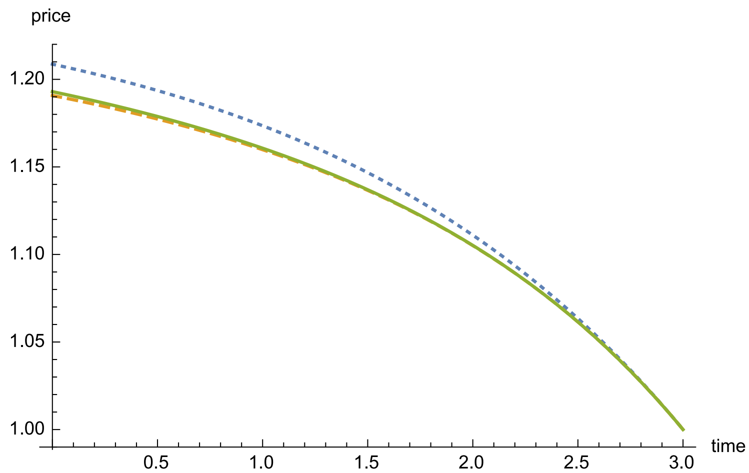


Figure 3.1: Equilibrium prices with both transaction and holding cost (solid), no transaction costs (dotted), and no holding costs (dashed).

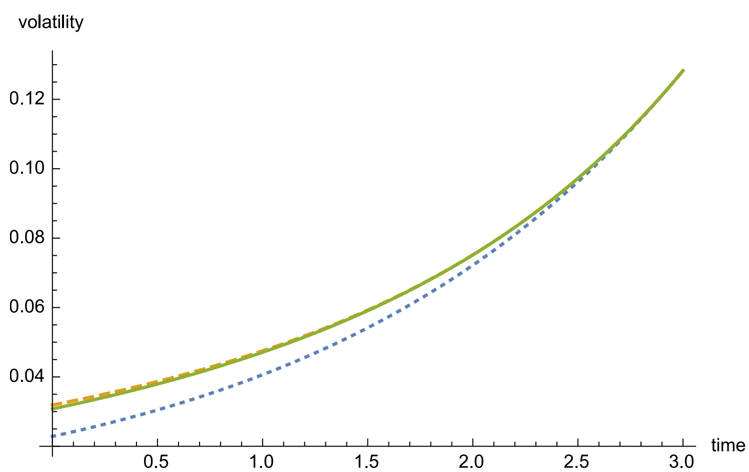


Figure 3.2: Equilibrium volatilities with both transaction and holding cost (solid), no transaction costs (dotted), and no holding costs (dashed).

Figures 3.1 and 3.2 also clearly show that the equilibrium price with holding and transaction costs is much closer to the risk-neutral price than to the frictionless one. This is not surprising, since $\gamma/\lambda = 0.1$ in this example. Accordingly, even though the qualitative predictions of both small-cost expansions are correct, only the small-holding-cost expansion provides useful quantitative approximations here. As shown in Figure 3.3, using the first-order correction term $v^0 + \gamma v^* - v^\gamma$ reduces the already small approximation error $v^0 - v^\gamma$ by another order of magnitude.

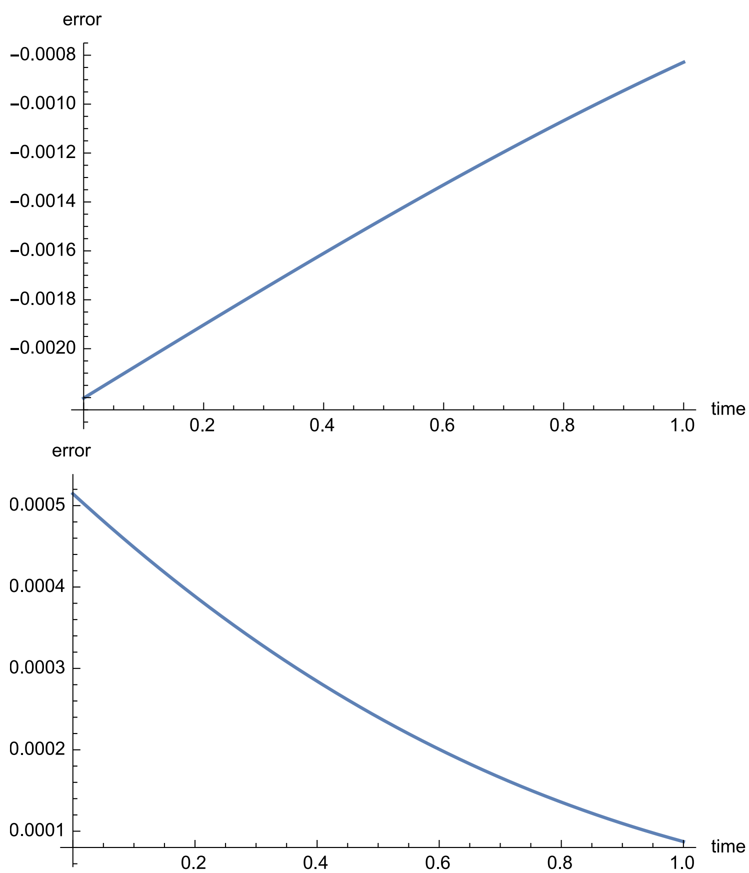


Figure 3.3: Approximation errors $v^0 - v^\gamma$ (top panel) and $v^0 + \gamma v^* - v^\gamma$ (lower panel).

Bibliography

[1] R. Almgren. Optimal trading with stochastic liquidity and volatility. *SIAM J. Financ. Math.*, 3(1):163–181, 2012. ↑113

[2] R.F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *J. Risk*, 3(2):5–39, 2001. ↑135

[3] Y. Amihud and H. Mendelson. Asset pricing and the bid-ask spread. *J. Financ. Econ.*, 17(2):223–249, 1986. ↑142

[4] S. Ankirchner and T. Kruse. Optimal position targeting with stochastic linear–quadratic costs. *Banach Center Publ.*, 104(1):9–24, 2015. ↑113, ↑116

[5] B. C. Arnold, C. A. Coelho, and F. J. Marques. The distribution of the product of powers of independent uniform random variables – A simple but useful tool to address and better understand the structure of some distributions. *J. Multivar. Anal*, 113:19–36, 2013. ↑101

[6] P. Bank, H.M. Soner, and M. Voß. Hedging with temporary price impact. *Math. Financ. Econ.*, 11(2):215–239, 2017. ↑113, ↑116

[7] M. Bardi. Explicit solutions of some linear-quadratic mean field games. *Netw. Heterog. Media*, 7(2):243–261, 2012. ↑71

[8] D. Becherer and M. Schweizer. Classical solutions to reaction-diffusion systems for hedging problems with interacting Itô and point processes. *Ann. Appl. Probab.*, 15(2):1111–1144, 2005. ↑110, ↑119, ↑120, ↑128

[9] M. Beiglböck, A. M. G. Cox, and M. Huesmann. The geometry of multi-marginal Skorokhod embedding. *Preprint arXiv:1705.09505v1*, 2017. ↑7

[10] M. Beiglböck, A. M. G. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math.*, 208(2):327–400, 2017. ↑7

[11] M. Beiglböck, A. M. G. Cox, M. Huesmann, N. Perkowski, and D. J. Prömel. Pathwise super-replication via Vovk's outer measure. *Finance Stoch.*, 21(4):1141–1166, 2017. ↑7

[12] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009. ↑14

[13] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013. ↑7

[14] M. Beiglböck, P. Henry-Labordère, and N. Touzi. Monotone martingale transport plans and Skorokhod embedding. *Stochastic Process. Appl.*, 127(9):3005–3013, 2017. ↑7

[15] M. Beiglböck, M. Huesmann, and F. Stebegg. Root to Kellerer. In *Séminaire de probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 1–12. Springer, Berlin, 2016. ↑7

[16] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016. ↑1, ↑3, ↑4, ↑5, ↑7, ↑10, ↑12, ↑39, ↑42, ↑44, ↑47, ↑50, ↑52, ↑62, ↑64

[17] M. Beiglböck and N. Juillet. Shadow couplings. *Preprint arXiv:1609.03340v1*, 2016. ↑43

[18] M. Beiglböck, M. Nutz, and N. Touzi. Complete duality for martingale optimal transport on the line. *Ann. Probab.*, 45(5):3038–3074, 2017. ↑2, ↑6, ↑7, ↑10, ↑11, ↑12, ↑21, ↑22, ↑25, ↑33, ↑37, ↑38, ↑39

[19] A. Bensoussan, J. Frehse, and S. C. P. Yam. *Mean field games and mean field type control theory.* Springer Briefs in Mathematics. Springer, New York, 2013. ↑70

[20] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control. The Discrete-Time Case.* Academic Press, New York, 1978. ↑37

[21] C. Bertucci. Optimal stopping in mean field games, an obstacle problem approach. *Preprint arXiv:1704.06553v1*, 2017. ↑72

[22] P. Besala. On solutions of Fourier's first problem for a system of non-linear parabolic equations in an unbounded domain. *Ann. Polon. Math.*, 13:247–265, 1963. ↑119

[23] H. S. Bhamra and R. Uppal. Asset prices with heterogeneity in preferences and beliefs. *Rev. Financ. Stud.*, 27(2):519–580, 2014. ↑113

[24] S. Biagini, B. Bouchard, C. Kardaras, and M. Nutz. Robust fundamental theorem for continuous processes. *Math. Finance*, 27(4):963–987, 2017. ↑7

[25] B. Bouchard, M. Fukasawa, M. Herdegen, and J. Muhle-Karbe. Equilibrium returns with transaction costs. *Finance Stoch.*, 22(3):569–601, 2018. ↑113, ↑116

[26] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.*, 25(2):823–859, 2015. ↑7

[27] A. Briani and P. Cardaliaguet. Stable solutions in potential mean field game systems. *NoDEA Nonlinear Differential Equations Appl.*, 25(1), 2018. ↑72

[28] J. I. Bulow, J. D. Geanakoplos, and P. D. Klemperer. Multimarket oligopoly: Strategic substitutes and complements. *J. Polit. Econ.*, 93(3):488–511, 1985. ↑72

[29] A. Buss, R. Uppal, and G. Vilkov. Asset prices in general equilibrium with transactions costs and recursive utility. *Preprint SSRN:2366296*, 2015. ↑113

[30] L. Campi and M. Fischer. N-player games and mean field games with absorption. *To appear in Ann. Appl. Probab.*, 2018. ↑71

[31] P. Cardaliaguet. Notes on mean field games. *(from P.-L. Lions' lectures at Collège de France)*, 2013. ↑70

[32] P. Cardaliaguet, F. Delarue, J. M. Lasry, and P. L. Lions. The master equation and the convergence problem in mean field games. *Preprint arXiv:1509.0250v1*, 2015. ↑71

[33] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I.* Springer, 2017. ↑70

[34] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications II.* Springer, 2017. ↑70

[35] R. Carmona, F. Delarue, and D. Lacker. Mean field games of timing and models for bank runs. *Appl. Math. Optim.*, 76(1):217–260, 2017. ↑71, ↑72, ↑84

[36] Á. Cartea and S. Jaimungal. A closed-form execution strategy to target volume weighted average price. *SIAM J. Financial Math.*, 7(1):760–785, 2016. ↑113, ↑116, ↑144

[37] P. Casgrain and S. Jaimungal. Mean-field games with differing beliefs for algorithmic trading. *Preprint arXiv:1810.06101v1*, 2018. ↑113

[38] A. Cecchin and M. Fischer. Probabilistic approach to finite state mean field games. *To appear in Appl. Math. Optim.*, 2018. ↑71

[39] A. Cecchin, P. Dai Pra, M. Fischer, and G. Pelino. On the convergence problem in mean field games: a two state model without uniqueness. *Preprint arXiv:1810.05492v1*, 2018. ↑71

[40] J.-H. Choi, K. Larsen, and D.J. Seppi. Smart TWAP trading in continuous-time equilibria. *Preprint arXiv:1803.08336v1*, 2018. ↑113

[41] A. M. G. Cox and J. Obłój. Robust pricing and hedging of double no-touch options. *Finance Stoch.*, 15(3):573–605, 2011. ↑7

[42] A. Danilova and C. Julliard. Understanding volatility, liquidity, and the Tobin tax. *Preprint, London School of Economics and Political Science*, 2019. ↑112, ↑141

[43] F. Delarue and R. Foguen Tchuendom. Selection of equilibria in a linear quadratic mean-field game. *Preprint arXiv:1808.09137v1*, 2018. ↑71

[44] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300:463–520, 1994. ↑34

[45] J. Detemple and S. Murthy. Intertemporal asset pricing with heterogeneous beliefs. *J. Econ. Theory*, 62(2):294–320, 1994. ↑113

[46] D. W. Diamond and P H. Dybvig. Bank runs, deposit insurance, and liquidity. *J. Polit. Econ.*, 91(3):401–419, 1983. ↑72

[47] Y. Dolinsky and H. M. Soner. Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Related Fields*, 160(1–2):391–427, 2014. ↑8

[48] Y. Dolinsky and H. M. Soner. Martingale optimal transport in the Skorokhod space. *Stochastic Process. Appl.*, 125(10):3893–3931, 2015. ↑8

[49] M. Fischer. On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab.*, 27(2):757–810, 2017. ↑71

[50] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. W. de Gruyter, Berlin, 3rd edition, 2011. ↑9

[51] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, N.J., 1964. ↑119, ↑122

[52] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1):312–336, 2014. ↑8

[53] N. Garleanu and L. H. Pedersen. Dynamic trading with predictable returns and transaction costs. *J. Finance*, 68(6):2309–2340, 2013. ↑113, ↑116, ↑144

[54] N. Garleanu and L. H. Pedersen. Dynamic portfolio choice with frictions. *J. Econ. Theory*, 165:487–516, 2016. ↑113, ↑116

[55] N. Ghoussoub, Y.-H. Kim, and T. Lim. Structure of optimal martingale transport in general dimensions. *Ann. Probab.*, 47(1):109–164, 2019. ↑7

[56] G. Guo, X. Tan, and N. Touzi. Optimal Skorokhod embedding under finitely many marginal constraints. *SIAM J. Control Optim.*, 54(4):2174–2201, 2016. ↑7

[57] H. Hau. The role of transaction costs for financial volatility: Evidence from the Paris bourse. *J. Eur. Econ. Assoc*, 4(4):862–890, 2006. ↑112, ↑141

[58] D. Heath and M. Schweizer. Martingales versus PDEs in finance: an equivalence result with examples. *J. Appl. Probab.*, 37(4):947–957, 2000. ↑124

[59] P. Henry-Labordère, J. Obłój, P. Spoida, and N. Touzi. Maximum maximum of martingales given marginals. *Ann. Appl. Probab.*, 26(1):1–44, 2016. ↑7

[60] P. Henry-Labordère, X. Tan, and N. Touzi. An explicit version of the one-dimensional Brenier's theorem with full marginals constraint. *Stochastic Process. Appl.*, 126(9):2800–2834, 2016. ↑8

[61] P. Henry-Labordère and N. Touzi. An explicit martingale version of the one-dimensional Brenier theorem. *Finance Stoch.*, 20(3):635–668, 2016. ↑7, ↑8

[62] M. Herdegen, J. Muhle-Karbe, and D. Possamaï. Equilibrium asset pricing with transaction costs. *Preprint arXiv:1901.10989v1*, 2019. ↑112, ↑113, ↑127, ↑141

[63] F. Hirsch, C. Profeta, B. Roynette, and M. Yor. *Peacocks and Associated Martingales, with Explicit Constructions*. Springer, Milan, 2011. ↑8

[64] D. Hobson. Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347, 1998. ↑7

[65] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, Berlin, 2011. ↑7

[66] D. Hobson and M. Klimmek. Robust price bounds for the forward starting straddle. *Finance Stoch.*, 19(1):189–214, 2015. ↑7

[67] D. Hobson and A. Neuberger. Robust bounds for forward start options. *Math. Finance*, 22(1):31–56, 2012. ↑7

[68] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007. ↑70

[69] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006. ↑70, ↑71

[70] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003. ↑9

[71] C. M. Jones and P. J. Seguin. Transaction costs and price volatility: Evidence from commission deregulation. *Am. Econ. Rev.*, 4(87):728–737, 1997. ↑112, ↑141

[72] E. Jouini and C. Napp. Consensus consumer and intertemporal asset pricing with heterogeneous beliefs. *Rev. Econ. Stud.*, 74(4):1149–1174, 2007. ↑113

[73] N. Juillet. Stability of the shadow projection and the left-curtain coupling. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1823–1843, 2016. ↑7

[74] N. Juillet. Martingales associated to Peacocks using the curtain coupling. *Electron. J. Probab.*, 23(8):1–29, 2018. ↑8

[75] S. Källblad, X. Tan, and N. Touzi. Optimal Skorokhod embedding given full marginals and Azéma–Yor peacocks. *Ann. Appl. Probab.*, 27(2):686–719, 2017. ↑8

[76] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, New York, 2nd edition, 1991. ↑128

[77] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997. ↑125

[78] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984. ↑2, ↑38

[79] M. Kohlmann and S. Tang. Global adapted solution of one-dimensional backward stochastic Riccati equations, with application to the mean–variance hedging. *Stochastic Process. Appl.*, 97(2):255–288, 2002. ↑113, ↑116

[80] D. Kramkov. Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probab. Theory Related Fields*, 105(4):459–479, 1996. ↑87

[81] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996. ↑122, ↑123

[82] D. Lacker. A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields*, 165(3-4):581–648, 2016. ↑71, ↑84

[83] D. Lacker. On the convergence of closed-loop Nash equilibria to the mean field game limit. *Preprint arXiv:1808.02745v1*, 2018. ↑71

[84] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006. ↑70, ↑71

[85] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006. ↑70

[86] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007. ↑70

[87] A. W. Lo, H. Mamaysky, and J. Wang. Asset prices and trading volume under fixed transaction costs. *J. Polit. Economy.*, 112(5):1054–1090, 2004. ↑113

[88] H. De March and N. Touzi. Irreducible convex paving for decomposition of multi-dimensional martingale transport plans. *Preprint arXiv:1702.08298v1*, 2017. ↑7

[89] L. Moreau, J. Muhle-Karbe, and H. M. Soner. Trading with small price impact. *Math. Finance*, 27(2):350–400, 2017. ↑135

[90] J. Muhle-Karbe and M. Nutz. A risk-neutral equilibrium leading to uncertain volatility pricing. *Finance Stoch.*, 22(2):281–295, 2018. ↑115

[91] V. N. Nair, L. A. Shepp, and M. J. Klass. On the number of crossings of empirical distribution functions. *Ann. Probab.*, 14(3):877–890, 07 1986. ↑98

[92] A. Neufeld and M. Nutz. Superreplication under volatility uncertainty for measurable claims. *Electron. J. Probab.*, 18(48):1–14, 2013. ↑8

[93] M. Nutz. A mean field game of optimal stopping. *SIAM J. Control Optim.*, 56(2):1206–1221, 2018. ↑72, ↑74, ↑75, ↑76, ↑80, ↑81

[94] M. Nutz and J. A. Scheinkman. Shorting in speculative markets. *Preprint SSRN:2969112*, 2017. ↑113, ↑125

[95] M. Nutz and F. Stebegg. Canonical supermartingale couplings. *Ann. Probab*, 46(6):3351–3398, 2018. ↑42

[96] M. Nutz and Y. Zhang. A mean field competition. *Preprint arXiv:1708.01308v1*, 2017. ↑71

[97] J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004. ↑7

[98] H. Robbins. A remark on Stirling's formula. *Am. Math. Mon*, 62(1):26–29, 1955. ↑102

[99] Y. Sannikov and A. Skrzypacz. Dynamic trading: price inertia and front-running. *Preprint SSRN:2882809*, 2016. ↑113

[100] J. Scheinkman and W. Xiong. Overconfidence and speculative bubbles. *J. Polit. Economy*, 111:1183–1219, 2003. ↑113

[101] J. Scheinkman and W. Xiong. Heterogeneous beliefs, speculation and trading in financial markets. In *Paris-Princeton Lectures on Mathematical Finance 2003*, volume 1847 of *Lecture Notes in Math.*, pages 217–250. Springer, Berlin, 2004. ↑110, ↑113

[102] H. M. Soner, N. Touzi, and J. Zhang. Dual formulation of second order target problems. *Ann. Appl. Probab.*, 23(1):308–347, 2013. ↑8

[103] F. Stebegg. Model-independent pricing of Asian options via optimal martingale transport. *Preprint arXiv:1412.1429v1*, 2014. ↑7

[104] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965. ↑1, ↑9

[105] D. W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley*

Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, pages 333–359. Univ. California Press, Berkeley, Calif., 1972. ↑114

[106] X. Tan and N. Touzi. Optimal transportation under controlled stochastic dynamics. *Ann. Probab*, 41(5):3201–3240, 2013. ↑8

[107] N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29 of *Fields Institute Monographs*. Springer, New York, 2013. ↑120, ↑121

[108] N. Touzi. Martingale inequalities, optimal martingale transport, and robust superhedging. In *Congrès SMAI 2013*, volume 45 of *ESAIM Proc. Surveys*, pages 32–47. EDP Sci., Les Ulis, 2014. ↑7

[109] S. R. Umlauf. Transaction taxes and the behavior of the Swedish stock market. *J. Financ. Econ.*, 2(33):227–240, 1993. ↑112, ↑141

[110] D. Vayanos. Transaction costs and asset prices: A dynamic equilibrium model. *Rev. Financ. Stud.*, 11(1):1–58, 1998. ↑113

[111] D. Vayanos and J.-L. Vila. Equilibrium interest rate and liquidity premium with transaction costs. *Econ. Theory*, 13(3):509–539, 1999. ↑113

[112] K. Weston. Existence of a Radner equilibrium in a model with transaction costs. *Math. Financ. Econ.*, 12(4):517–539, 2018. ↑113

[113] G. Žitković. A filtered version of the bipolar theorem of Brannath and Schachermayer. *J. Theoret. Probab.*, 15:41–61, 2002. ↑87